DOCUMENT RESUME

ED 048 462 VT 012 451

TITLE Longitudinal Evaluation Study of Five Manpower Training Programs: Investigation of the Study Sample Design.

INSTITUTION Operations Research, Inc., Silver Spring, Md.

SPONS AGENCY Office of Economic Opportunity, Washington, D.C. Office of Planning, Research, and Evaluation.

PUB DATE 30 Jun 69

NOTE 149p.

EDRS PRICE EDRS Price MF-$0.65 HC-$6.58

DESCRIPTORS Cost Effectiveness, Employment, Evaluation Methods, *Evaluation Techniques, *Federal Programs, Income, Longitudinal Studies, *Manpower Development, Participant Characteristics, *Program Evaluation, *Research Design, Research Methodology

ABSTRACT

This report develops the sample design for evaluating five Federal manpower programs in terms of their effects upon trainees' incomes. The study involves a large 10-city sample of trainees in Job Corps, Neighborhood Youth Corps, Job Opportunities in the Business Sector, and Manpower and Development Act Institutional Training Programs (New Careers was originally included, but later deleted because of its lack of similarity to the other programs). After interviews and tests at the beginning of training, the enrollee will be followed through training and for 12 months afterward. A control group will be interviewed at the same intervals. This report describes the sample design and estimates the necessary sample size for given levels of training. (Author/BH)

ED0 48462

0-2
2

OPERATIONS RESEARCH, Inc.

SILVER SPRING, MARYLAND

A Subsidiary of Leasco Systems & Research Corporation

VT012451

ED0 48462

OPERATIONS RESEARCH, Inc.
SILVER SPRING, MARYLAND

LONGITUDINAL EVALUATION STUDY OF FIVE MANPOWER TRAINING PROGRAMS: INVESTIGATION OF THE STUDY SAMPLE DESIGN

30 June 1969

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Prepared under OEO Contract No. B99-4783
for the Office of Economic Opportunity
and the U. S. Department of Labor
Washington, D. C.

ACKNOWLEDGEMENT

This investigation was aided by a large number of individuals in OEO and DOL. Advice and assistance was freely given by staffs of the programs to be evaluated in the study. Special tabulations of program data were furnished by the Office of Manpower Management Data Systems, DOL. Particular acknowledgement is due Mrs. Lillian Regelson and Gilmore P. Wheeler of the Division of Research Planning, Programming, and Evaluation, OEO; and Stan D. Markuson, Dr. William Tash, and Mrs. Gail Osborne of the Office of Evaluation Studies, Manpower Administration, DOL.

SUMMARY

PURPOSE

1. The purpose of this report is to document the analyses carried out for the design of a sample for the longitudinal evaluation study of five major U.S. Government manpower training programs for the Office of Economic Opportunity (OEO) and the Department of Labor (DOL).

SCOPE

2. The general design of the longitudinal evaluation study may be described as follows. A set of areas will be selected and, within these, longitudinal data will be collected on enrollees in the programs of interest and on a group of selected individuals who are not program enrollees, which will serve as a control group. Ancillary data concerning the selected areas that are relevant for the study analyses will also be collected. The manpower training programs considered are:

a. Job Opportunities in the Business Sector (JOBS - Contract component)

b. Job Corps

c. Manpower Development and Training Act (MDTA - Institutional component)

d. Neighborhood Youth Corps (NYC - Out-of-School component)

e. New Careers (NC).

Program enrollees will be interviewed:

a. At the time of their enrollment to obtain detailed information about their background and characteristics (pre-program interview)

b. At the time of leaving the program, either by completion of their training or by dropping out (post-program interview)

c. Thereafter, at times:

    1. 3 months after leaving the program

    2. 9 months after leaving the program

    3. 18 months after leaving the program

    to collect data on their subsequent employment experience and income.

The study will also collect relevant data on the kind, extent, and quality of training and other services enrollees receive. Individuals in the control group will be interviewed at corresponding points in time.

3. This report covers the selection of a set of study areas and a preliminary discussion of the sampling of program enrollees and matching control cases to be followed in the study within the selected areas.

APPROACH

4. OEO and DOL had established preliminary specifications for a study sample of 10 areas and 10,000 study persons, and had requested recommendations from the study group as to the adequacy of these specifications. This request was a major focus of the work summarized in this report.

5. The approach followed in the investigation was to:

a. Establish a reasonably efficient sample design, considering the major—and possibly conflicting—objectives of the study, within the OEO-DOL specifications.

b. Develop rough guides as to sampling errors to be expected with the sample design for some major estimates to be produced by the study, as a basis for review of the adequacy of the initial specifications and possible desirable revisions.

6. The type of design investigated may be described in a general way as a two-stage sample, the two stages being:

a. A sample of areas

b. Within each of the selected areas, a sample of enrollees from each of the study programs and a sample of matching control cases.

The details of the design to be determined are:

a. The definition of the universe of areas to be sampled for the study

b. The method of selection of the areas

c. The numbers of program enrollees to be selected and the method of selection

d. The specific definition of the control populations for the study, within the general recommendation made to OEO and DOL by LS&R that the control cases be samples of the program target populations

e. The numbers of control cases to be selected and the method of selection

f. The estimation techniques to be used for preparing estimates from the study data, for purposes of developing estimates of sampling errors.

7. The analyses were carried out with a view to programmatic uses of the data to come from the study, rather than just a benefit-cost analysis for the particular enrollees in the programs at the time of the study. Programmatic uses of the data are considered to primarily require estimates of program impacts, and of the sampling errors of such estimates, rather than tests of significance.

FINDINGS

Preliminary OEO-DOL Specifications

8. The preliminary specifications for the study sample established by OEO and DOL can be expected to provide estimates and analyses within programmatically useful limits of sampling error for four programs:

Job Corps
JOBS (Contract)
MDTA (Inst.)
NYC (O/S).

The criterion of programmatically useful sampling errors adopted in this report is that the uncertainty in estimated post-program changes in annual earnings be within \$300-\$400. Put another way, there should be high assurance that if changes of this order of magnitude exist they will be detected statistically in the study. It is considered that changes on the order of \$50-\$100 a year are of questionable significance either for program planning or for potential program enrollees. Moreover, there is not sufficient accuracy in techniques for measuring total annual income to feel much confidence in differences on the order of \$50-\$100 a year.

Speculated Sampling Errors of Major Estimates

9. Post-Program Changes in Income. Estimates of sampling errors to be expected in estimating post-program changes in average annual income per program enrollee, compared with their controls, were constructed using data from

earlier evaluation studies made available by OEO and DOL. Because of the limitations in the available data, these estimates of expected sampling errors are referred to in the report as "speculated sampling errors." Despite the qualifications and uncertainties that are attached to the speculated sampling errors, it is felt that they provide a reasonable indication of the level of sampling errors to be expected in independent estimates from the study. An illustrative summary of the results of the sampling error analysis is given in the following tabulation for estimates of post-program change in average annual earnings for all enrollees in a given program compared with their controls.

If the average annual earnings per control case is	The chances are about 2 out of 3 that the difference between the estimated change and the change in annual earnings derived from a study based on all enrollees in the program would be less than	
	Estimated change for study areas	National projection of change to all areas
$ 500	$ 40-$ 50	$ 60-$ 75
$1,000	$ 65-$ 75	$110-$110
$2,000	$110-$115	$150-$160

The chances are about 19 in 20 that the difference between the change estimated from the study sample and that which would be found from a study of all enrollees would be less than twice the limits given in this tabulation. In Section III of this report a more detailed analysis of the speculated sampling errors for estimates of post-program changes in earnings by age-sex-race group, and for estimated benefit-cost ratios, is presented. Also presented are estimates of true changes in post-program earnings for which the odds (or probability) that the change would be detected statistically in the study are suitably high.

10. Benefit-Cost Ratios. It is speculated that the sampling errors of benefit-cost ratios estimated from the study might be as high as 10 to 20 percent for ratios estimated for the study areas, and 15 to 25 percent for national projections of ratios to all areas, under some combination of benefits and costs.

11. Although the speculated sampling errors appear to be high compared to the specifications for precision ordinarily met in survey studies, it is suggested that they are useful for purposes of the study analysis. For example, a 20 percent sampling error in observed benefit-cost ratios would have the implications summarized in the following tabulation.

If the observed benefit-cost ratio is	The chances are about 2 out of 3 that the difference between the observed ratio and the benefit-cost ratio from a study of all program enrollees would be less than
1	0.2
2	0.4
5	1.0
10	2.0

The uncertainty due to sampling which is illustrated by this tabulation is relatively small compared to that arising from other sources of uncertainty which affect the estimated benefit-cost ratios. Among such factors which affect the level of an estimated benefit-cost ratio are the assumptions as to the patterns of benefits to be projected for time periods not directly observed, the length of the time horizon over which benefits are projected, and the choice of an appropriate rate for discounting future benefits,

RECOMMENDATIONS FOR SAMPLE DESIGN

Program Coverage

12. It is recommended that the New Careers program be dropped from the study or be budgeted separately. The area distribution of New Careers is markedly different than that of the other programs. To include it without separate budgeting, and correspondingly reduce the sample size for the other programs, would increase the sampling errors of estimates and analyses for the other programs without providing estimates for New Careers on a basis comparable to those for the other programs. If the New Careers program is budgeted separately, some additional areas should be selected for the study to strengthen the estimates and analyses for New Careers. Also, a longer initial period for accumulating the desired sample of enrollees should be planned for than in the case of the other programs.

Universe of Study Areas

13. It is recommended that the universe of areas from which the study areas will be selected be taken as the Labor Market Areas corresponding to SMSAs of 500,000 or more population in 1960 with central city of 250,000 or more. This universe includes 43 of the 46 JOBS cities in conterminous U.S. The JOBS cities not included are El Paso, Omaha, and Tulsa. Excluded outside conterminous U.S. are Honolulu and three JOBS cities in Puerto Rico. The establishing of a universe of study areas for sampling helps clarify the universe for which statistically-based inferences from the study will be possible; and the universe for which, since it was not sampled, inference will depend on subject-matter expertise. From the definition of the universe of study areas, it is clear that the study will be an urban, not rural, one. Such rural places as may be represented in the study sample will be of the type found in the SMSAs of large cities.

Sample of Areas

14. It is recommended that the 10 areas in which the study will be conducted be selected as a probability sample. Since individuals will be designated for the study on enrollment, as they are referred by the programs, the fact that they are the subjects of evaluation will be known to program staffs. This fact could be reflected in special selection and/or treatment of enrollees. If favorable outcomes for enrollees are observed in a probability sample of areas, the inference is that what was accomplished in the study areas can be accomplished elsewhere. If the study areas were chosen subjectively it would be difficult to disprove the argument that this inference should not be drawn.

15. A recommended design for the sampling of areas is described. The design is intended to be reasonably efficient for the various study objectives and programs.

Sample of Individuals

16. It is recommended that the initial sample of individuals to be included in the study consist of 7,500 program enrollees and 3,500 control cases, to be allocated equally by program group and by area to the extent feasible. The cost of this sample of 11,000 individuals is believed to be equivalent to that intended under the preliminary OEO-DOL specification for 10,000 individuals to be included in the study. It would provide a total initial sample of 1,250 enrollees in each of the Job Corps and NYC (O/S) programs, and in each of the two age groups (under 22, and 22 and over) of the JOBS and MDTA (Inst.) programs; and 1,750 control cases for each of the two age groups. The larger size of the control groups relative to the program groups is to provide for attrition of the control sample when individuals selected as controls enroll in programs, as they are expected to over the life of the study. If the avoidable sample attrition is held to 20 percent over the life of the study, this would provide a final sample of 1,000 individuals per group for the analysis. This sample allocation can be expected to permit estimates of post-program changes in average annual earnings per enrollee, and benefit-cost ratios, within programmatically useful limits of sampling error.

17. It is recommended that the sample of enrollees for each program be allocated equally among the four race-sex groups to the extent feasible. This allocation is intended to provide estimates of post-program changes in average annual earnings per enrollee by race-sex group within useful limits of sampling error for each, as well as to provide a basis for further analyses of the impacts of program components and the factors in success or failure for each of the age-sex groups. From the data available for analysis on this point, it appears that an allocation of the sample by race and sex proportional to the mix of program enrollees that might be encountered during the study sampling period would be likely to provide poor estimates for whites and, to a lesser extent, for females.

Sampling of Program Enrollees and Control Cases

18. Sampling Approach. The study, being based on a sample of enrollees entering the programs during a particular period of time (and a matching sample of control cases) is, in fact,a study of a particular cohort. Generalization to other populations and different program designs depends upon subject matter expertise to the extent to which the processes studied are not stationary or stable. Such generalization is aided by careful specification and description of the study cohort itself, so that differences between the cohort and other populations of interest can be adequately understood.

19. It is recommended that the program enrollees and control cases for the study be selected as probability samples. In the report,some preliminary comments are made on the principles to be followed in sampling of program enrollees and control cases in the study. It is recommended that the study sampling be extended over as long a time period as feasible. This enlarges the size of the cohort, and helps avoid the possibility of being forced to accept an undesirable distribution of enrollees by characteristics.

20. Some Issues for Decision. Several issues for decision are identified in the report:

a. Programs enroll some individuals who would be ineligible under a strict application of program criteria. It is tentatively proposed to select the enrollee sample only from those individuals who meet program eligibility criteria. This is subject to review after the New York City pretest.

b. Questions have been raised as to whether the enrollee population to be sampled for the study should be restricted, for example by excluding older or disabled persons. These are to be resolved by consideration of the benefits for the study to be gained by the restrictions.

c. It is tentatively proposed in each study area, to sample program activities over the entire SMSA. This is subject to review on the basis of the expected impact on program costs.

Some Opportunities in the Research Design

21. The limitations of the sample design are, first of all, inherent in the research design itself and the fact that the study will be based on observation rather than experimentation. Some opportunities in the research design for strengthening the study data and the span of inference from the study if funds are available are identified in the report.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENT i
SUMMARY iii
LIST OF FIGURES xv
LIST OF TABLES xvii
I. INTRODUCTION 1
    OBJECTIVE 1
    SCOPE 1
    APPROACH 2
    LIMITATIONS 3
    ORGANIZATION OF REPORT 4
II. THE UNIVERSE OF AREAS FOR THE STUDY 5
    INTRODUCTION 5
    DEFINITION OF "AREA" FOR THE STUDY 5
    POTENTIAL UNIVERSE OF STUDY AREAS 5
        Large Urban Areas and Associated Labor Markets; Further Analysis of Potential Universe
    RECOMMENDED UNIVERSE OF STUDY AREAS 8

III. DEVELOPMENT OF SAMPLE DESIGN 15

INTRODUCTION 15

DESIGN APPROACH 15

Study Objectives for Design, Study Estimates for Design; Implications of Study Objectives for Design; Estimation Approaches; Design Approach

DESIGN INVESTIGATION 19

Approach; Optimization Analysis; Considerations for More Than One Program: Summary of Investigation of Sampling the Study Areas; Recommended Design for Sample of Areas

SPECULATED SAMPLING ERRORS 33

Approach; The Speculated Sampling Errors; Application of Speculated Sampling Errors; Applications for Benefit-Cost Ratios; Implications for Benefit-Cost Ratios

SAMPLE ALLOCATION WITH A FIXED TOTAL COST 51

Study Sample Excluding New Careers Program; New Careers Program

IV. A 10-AREA SAMPLE FOR THE STUDY 55

INTRODUCTION 55

STRATIFICATION VARIABLES 55

Program Evaluation Data; Political and Social Stresses; Socio-Economic Data

DESCRIPTION OF THE STRATA 60

Note on the Stratification, Characteristics of the Strata

AREAS AND THEIR SELECTION 61

V. SAMPLING PROGRAM ENROLLEES AND CONTROL CASES 67

INTRODUCTION 67

SAMPLING OF PROGRAM ENROLLEES 68

Sampling of Projects; Sampling of Enrollees Within Projects

SAMPLING OF CONTROLS 69

Approach; Scheduling of Screening; Matching in Selection of Controls

SOME POLICY ISSUES 70

Enrollee Population to Be Sampled; Target Population Sampling

VI. SOME OPPORTUNITIES IN THE RESEARCH DESIGN 73

INTRODUCTION 73

PROGRAM COVERAGE 73

PROGRAM ORGANIZATION 73

TARGET POPULATION 74

VII. CONCLUSIONS AND RECOMMENDATIONS 75

PRELIMINARY OEO-DOL SAMPLE SPECIFICATIONS 75

UNIVERSE OF STUDY AREAS 76

STUDY SAMPLE DESIGN 76

Sample of Areas; Allocation of Sample of Individuals

SPECULATED SAMPLING ERRORS 77

Post-Program Changes in Income; Benefit-Cost Ratios

SAMPLING OF PROGRAM ENROLLEES AND CONTROL CASES 79

Comments on the Sampling Approach; Some Issues for Decision

SOME OPPORTUNITIES IN THE RESEARCH DESIGN . . 80

APPENDIX A: TABLES FOR SECTION II 81

APPENDIX B: BIBLIOGRAPHIC NOTE FOR SECTION III . . . 87

APPENDIX C: REGRESSION ANALYSIS FOR ESTIMATES BY ANALYSIS GROUPS 89

APPENDIX D: SAMPLE DESIGN ANALYSIS 95

APPENDIX E: NUMERICAL ESTIMATES OF VARIANCE AND COST PARAMETERS 111

APPENDIX F: MATCHING OF CONTROL CASES WITH PROGRAM ENROLLEES 131

LIST OF FIGURES

Page

1. Speculated Standard Error of Estimated Average Annual Post-Program Earnings per Terminee, Fixed Set of Areas, by Level of Earnings 39

2. Speculated Standard Error of National Projection of Average Annual Post-Program Earnings per Terminee, 10-Area Design, by Level of Earnings 40

E.1 Generalized Within Component of Relvariance per Person, Average Annual Post-Program Earnings: Terminees and Controls 116

E.2 Relationships of Generalized Between and Within Components of Variance, Average Annual Earnings . . . 118

LIST OF TABLES

Page

1. Number of SMSAs by 1960 Population of Central City and SMSA, for all SMSAs and for JOBS Cities and DOL Urban Centers With Manpower Projects Active, as of February 1960 7

2. Federal Funds Authorized for Active Projects in All DOL Manpower Programs and for Four Study Programs, by Type of Program Area, as of February 1969 9

3. Positions Authorized for DOL Urban Centers in Four Study Manpower Programs, by Program and Type of Program Area, as of February 1969 11

4. Enrollment in DOL Urban Centers in Four Study Manpower Programs, by Program and Type of Program Area, as of February 1969 13

5. Illustration of Program Estimates for Sample Design Analysis 21

6. Illustration of Program Comparisons for Sample Design Analysis 27

7. Speculated Sampling Error of Average Annual Post-Program Earnings per Terminee, Fixed Set of Areas and National Projection, by Number of Terminees in Sample 36

8. Speculated Sampling Error of Average Annual Earnings per Control Case, Fixed Set of Areas and National Projection, by Number of Control Cases in Sample . . . 37

9. Speculated Sampling Error of Estimated Average Increase in Annual Post-Program Earnings per Terminee, Fixed Set of Areas 41

10. Speculated Sampling Error of National Projection of Average Increase in Annual Post-Program Earnings per Terminee 44

11. Least Significant Increases in Program Earnings 48

12. Increases in Post-Program Earnings That Will Be Detected with Specified Probabilities 49

13. Socio-Economic Data for Study Universe SMSAs 58

14. Strata for 10-Area Sample 62

15. 10-Area Sample Program Status, as of February 1969 65

A.1 50 JOBS Cities 82

A.2 Regional Offices for HEW, LABOR, OEO, SBA 84

A.3 DOL Manpower Programs Reported in Status Report of Projects Active as of February 1969 85

E.1 Characteristics of Terminees From NYC Out-of-School Programs in Sites Used for Analysis 121

E.2 Characteristics of Terminees From Job Corps in Cities of 1,000,000 or More Population Used for Analysis 122

E.3 Characteristics of Terminees From Job Corps in Cities of 250,000 - 999,000 Population Used for Analysis 124

E.4 Relvariance per Terminee of Annual Post-Program Earnings Between NYC Out-of-School Terminees Within Site 126

E.5 Relvariance per Terminee of Annual Earnings Between Job Corps Terminees Within Region, by Size of Place of Interview 127

E.6 Relvariance per Terminee in Average Annual Post-Program Earnings, Between Terminee Within Site, NYC Out-of-School Program 128

E.7 Differences Between Coefficients of Variation for Average Annual Earnings, as Estimated From Survey Data and Derived From Generalized Curves 129

E.8 Estimated Average Relvariances per Site of Earnings Per Week Between Sites Within Stratum, NYC Out-of-School Terminee Survey Data 130

I. INTRODUCTION

OBJECTIVE

1.1 The objective of this report is to document the analyses carried out for the design of a sample for the longitudinal evaluation study of five major U.S. Government manpower training programs for the Office of Economic Opportunity (OEO) and the Department of Labor (DOL).

SCOPE

1.2 The general design of the longitudinal evaluation study may be described as follows. A set of areas will be selected and, within these, longitudinal data will be collected on enrollees in the programs of interest and on a group of selected individuals who are not program enrollees, which will serve as a control group. Ancillary data concerning the selected areas that are relevant for the study analyses will also be collected.

1.3 The manpower training programs included in the study are:

a. Job Opportunities in the Business Sector (JOBS-Contract component)

b. Job Corps

c. Manpower Development and Training Act (MDTA-Institutional component)

d. Neighborhood Youth Corps (NYC-Out-of-School component)

e. New Careers (NC).

Program enrollees will be interviewed:

- a. At the time of their enrollment to obtain detailed information about their background and characteristics (pre-program interview)
- b. At the time of leaving the program, either by completion of their training or by dropping out (post-program interview)
- c. Thereafter, at times:
  1. 3 months after leaving the program
  2. 9 months after leaving the program
  3. 18 months after leaving the program

  to collect data on their subsequent employment experience and income.

The study will also collect relevant data on the kind, extent, and quality of training and other services enrollees receive. Individuals in the control group will be interviewed at corresponding points in time.

1.4 This report covers the selection of a set of study areas and a preliminary discussion of the sampling of program enrollees and matching control cases to be followed in the study within the selected areas.

APPROACH

1.5 OEO and DOL had established preliminary specifications for a study sample of 10 areas and 10,000 study persons, and had requested recommendations from the study group as to the adequacy of these specifications. This request was a major focus of the work summarized in this report.

1.6 The approach followed was to:

- a. Establish a reasonably efficient sample design, considering the major—and possibly conflicting—objectives of the study, within the OEO-DOL specifications.
- b. Develop rough guides as to sampling errors to be expected with the sample design for some major estimates to be produced by the study, as a basis for review of the adequacy of the initial specifications and possible revisions desirable.

1.7 The type of design investigated may be described in a general way as a two-stage sample, the two stages being:

- a. A sample of areas
- b. Within each of the selected areas, a sample of enrollees from each of the study programs and a sample of matching control cases.

The details of the design to be determined are:

a. The definition of the universe of areas to be sampled for the study

b. The method of selection of the areas

c. The numbers of program enrollees to be selected and the method of selection

d. The specific definition of the control populations for the study, within the general recommendation made to OEO and DOL by LS&R that the control cases be samples of the program target populations

e. The numbers of control cases to be selected and the method of selection

f. The estimation techniques to be used for preparing estimates from the study data, for purposes of developing estimates of sampling errors.

LIMITATIONS

1.8 The limitations of the sample design are, first of all, those inherent in the research design itself. A theoretically ideal research design does not seem to be definable in the sense that the potential target populations, program operations, and area environments to which it would be desirable to generalize the study findings are not well-defined. Even treating all of these as well-defined and even finite populations, generalization to combinations of populations, programs, and environments not sampled must necessarily have a large component of subject-matter expertise and judgment.

1.9 Even for the sampled universe, a theoretically ideal research design does not seem to be achievable given the administrative constraints on the introduction of randomization in the study design posed by the ongoing programs. In an idealized research design, both the program enrollees and the control cases with which they are compared to evaluate program impacts should represent probability samples of the same populations. Some kind of random selection of individuals in the target populations of the programs to be enrolled in the programs and serve as the enrollee sample would then be required. Since this is not possible, the study will be based on observation rather than experimentation. Finally, on the enrollee side, the identity of the program enrollees in the study will be known to program staffs. On the control group side, it will not be possible to prevent control cases from enrolling in programs as the study continues.

1.10 Within these limitations, the effort in the sample design is to create as firm a statistical base as possible for the analysis and interpretation of the findings to come from the study. However, the depth to which it was feasible to carry the investigation is limited by the availability of relevant data. It is exceedingly fortunate that OEO was able to arrange access for the study to data from a national study of Job Corps terminees, and DOL to data from a national study of NYC out-of-school terminees. However, even with these data, it is necessary to also depend in part on informed judgment and speculation to develop a recommended design.

ORGANIZATION OF REPORT

1.11 In Section II, the proposed universe of study areas is defined. The development of a proposed sample design is described in Section III, and estimates are given for sampling errors expected for estimating post-program earnings of program participants and for program-control and program-program comparisons in the study. Section IV describes a 10-area sample for the study. The sampling of program enrollees and controls within selected areas is discussed in Section V. Some opportunities in the research design open to the study over time are noted in Section VI. The conclusions and recommendations from the investigation are summarized in Section VII. Detailed supporting material appears in Appendices A through F.

II. THE UNIVERSE OF AREAS FOR THE STUDY

INTRODUCTION

2.1 The establishing of a universe of study areas for sampling clarifies the universe for which statistically-based inferences from the study will be possible and the universe for which, since it was not sampled, inference will depend on subject-matter expertise. In this section of the report,a definition of the universe of areas is proposed for the study.

DEFINITION OF "AREA" FOR THE STUDY

2.2 Because of the area-related nature of manpower problems and effective solutions for them, the United States is viewed as the sum of its component areas. Specifically, the achievements and benefits of the manpower training programs will be the sum of the achievements and benefits area by area, the programs in each area being adapted to local needs and conditions. Definitions for such areas should reflect the geographic domain of program units and the socio-political units with which they interact. Carrying this a step further, it seems reasonable to define the smallest area unit in terms of a combination of the concepts of "labor market," as defined by the Department of Labor (DOL) and political units. These areas are the first-stage sampling units of the study.

POTENTIAL UNIVERSE OF STUDY AREAS

Large Urban Areas and Associated Labor Markets

2.3 Again,because of the area-related nature of the study problem, it is desirable to be able to compare the five programs covered by the study within identical areas to the extent feasible. Since the JOBS program had been operating in only 50 cities, these were taken as the starting point for defining the study universe. Review of these cities showed that all were over 250,000 in population in 1960. This confirms that the study will relate to large urban areas and their associated labor markets.

Further Analysis of Potential Universe

2.4 Further review was then carried out to characterize the nature of these cities as a universe of potential areas for the study. First, the JOBS program cities include Honolulu, and this was excluded from the study on the a priori ground that the cost of operating the study there, if it were selected, could not be justified.[1] The remaining 49 cities were then compared with the 77 urban centers other than Honolulu in which DOL manpower programs were active as of February 1969.[2] The 49 JOBS cities are located in 46 Standard Metropolitan Statistical Areas (SMSAs). The 77 DOL urban centers are located in 69 SMSAs, except for Gainesville, Ga. (1960 population 16,523) and Eagle Pass, Tex. (1960 population 12,094) which are not located in SMSAs. A cross-tabulation of the SMSAs in which the 49 JOBS cities and 75 DOL urban centers are located, by size of central city and size of SMSA, is given in Table 1.[3]

2.5 Table 1 shows that cities of 250,000 or more population in 1960 are associated primarily with SMSAs of over 500,000 population. Only 4 of these 47 cities are associated with SMSAs of 250,000 to 500,000 population in 1960: El Paso (Tex.), Omaha (Neb.), Tulsa (Okla.), and Wichita (Kans.). El Paso, Omaha, and Tulsa are all both JOBS cities and DOL urban centers; Wichita is neither.

2.6 Table 2 shows that these cities and their SMSAs account for about 92 percent of all Federal funds authorized for the four study DOL manpower programs as of February 1969, and about 83 percent of the funds (including CEP funding) authorized for all DOL manpower programs. In terms of funds authorized for urban places, the corresponding coverage is 95 percent for the four study DOL manpower programs and 90 percent for all DOL manpower programs. Table 3 shows the distribution of positions in the four study manpower programs directly authorized under programs funds for DOL urban centers as of February 1969. As expected, this distribution shows concentration in the JOBS cities similar to that of funds authorized.

2.7 Enrollment in the four study programs can be financed through CEP funding as well as direct program funding. Statistics on enrollment through both sources of funding are available only for current enrollment in active projects. Table 4 shows the distribution of this enrollment as of February 1969, by program and funding source. Again, the distribution shows high concentration in the JOBS

---

[1] See Appendix A for a list of the 50 JOBS cities.

[2] U.S. Department of Labor, Manpower Administration, Status Report of Projects Active as of February 1969, 23 April 1969 (Processed).

[3] SMSAs are substantially the same as DOL's Labor Market Area (LMAs) and make a convenient geographic unit for which statistical data can be readily obtained.

TABLE 1

NUMBER OF SMSAs BY 1960 POPULATION OF CENTRAL CITY AND SMSA, FOR ALL SMSAs AND FOR JOBS CITIES AND DOL URBAN CENTERS WITH MANPOWER PROJECTS ACTIVE, AS OF FEBRUARY 1969[1]

1960 Population of Central City	1960 Population of SMSA[2]								
	500,000 or More			250,000 to 500,000			Under 250,000		
	All SMSAs	DOL Urban Centers	JOBS Cities	All SMSAs	DOL Urban Centers	JOBS Cities	All SMSAs	DOL Urban Centers	JOBS Cities
250,000 or more	43[3]	43[3]	43[3]	4[4]	3[5]	3[5]	—	—	—
Under 250,000	12	3	—	50	11	—	120	9	—
Total	55	46	43	54	14	3	120	9	—

[1] Centers with projects active as shown in U.S. Department of Labor, Manpower Administration, Status Report of Projects Active as of February 1969, 23 April 1969; data on SMSAs from U.S. Bureau of the Census, Metropolitan Area Statistics, September 1968 and County and City Data Book, April 1967.

[2] Excludes Honolulu (Hawaii) SMSA and the three SMSAs in Puerto Rico.

[3] The cities in each class are identical.

[4] El Paso (Tex.), Omaha (Neb.), Tulsa (Okla.), Wichita (Kans.).

[5] El Paso (Tex.), Omaha (Neb.), Tulsa (Okla.).

cities. It may also be observed that, apart from New Careers, CEPs are not a major source of funds in the study programs. For JOBS cities, enrollment through CEPs as a percent of total current enrollment is as follows:

Program	Percent
MDTA (Institutional)	8.4
JOBS	—
NYC (Out-of-School)	10.6
New Careers	71.4

2.8 Statistics are not currently available as to the numbers of Job Corps enrollees who enrolled from a JOBS city or SMSA, or returned to one after termination in the program.[4/] Therefore, the Job Corps has not been included in the preceding analysis. Some speculation suggests that on the order of 30,000 Job Corps enrollments a year may be associated with JOBS cities or their SMSAs. This would be a programmatically significant number.

RECOMMENDED UNIVERSE OF STUDY AREAS

2.9 The conclusion drawn from this review is that the cities of 250,000 or more population in 1960 and their SMSAs constitute a programmatically useful universe of potential areas for study. The JOBS cities are a very close approximation to this universe. A question may be raised, however, as to the inclusion of cities of 250,000 or more which were located in SMSAs of under 500,000 population in the U.S. From the point of view of a programmatically oriented study, therefore, it appears inefficient to include them in the reference universe. Accordingly, it is recommended that the reference universe be defined as the Labor Market Areas corresponding to the 43 SMSAs of 500,000 or more population with central city of 250,000 or more population in 1960. The central city of each of these SMSAs was both a DOL urban center with projects active as of February 1969 and a JOBS city, the SMSA containing additional program cities in some cases.

[4/] The Job Corps is undertaking the task of providing some data on these points for the study. These data will be incorporated in the study documentation when they are available.

TABLE 2
FEDERAL FUNDS AUTHORIZED FOR ACTIVE PROJE
FOUR STUDY PROGRAMS, BY TYPE OF

All DOL Manpower Programs (Millions of Dollars)[2]					
		JOBS Cities			
Program Region[4]	Total, All Centers	In SMSAs of 500,000 or More Population in 1960	In SMSAs of Under 500,000 Population in 1960	Other DOL Urban Centers	DOL Rural Centers
I.	53.6	24.4	—	29.2	—
II.	158.8	153.7	—	5.1	—
III.	143.5	111.2	—	6.2	26.1
IV.	81.6	41.5	—	20.0	20.1
V.	185.1	166.6	—	7.3	11.2
VI.	88.3	66.4	5.1	10.7	6.1
VII.	54.6	41.7	4.0	3.7	5.2
VIII.	165.6	155.3	—	6.9	3.4
Total, all regions	931.1	760.8	9.1	89.1	72.1
Percent	100.0	81.7	1.0	9.6	7.7

[1] U.S. Department of Labor, Manpower Administration, Status Report of

[2] See Appendix A for list of manpower programs.

[3] MDTA - Institutional component, JOBS - Contract component, Neighborh
New Careers.

[4] See Appendix A for definition of Program Regions.

TABLE 2

VE PROJECTS IN ALL DOL MANPOWER PROGRAMS AND FOR
TYPE OF PROGRAM AREA, AS OF FEBRUARY 1969[1]

	Four Study DOL Manpower Programs (Millions of Dollars)[3]				
		JOBS Cities			
DOL Rural Centers	Total, All Centers	In SMSAs of 500,000 or More Population in 1960	In SMSAs of Under 500,000 Population in 1960	Other DOL Urban Centers	DOL Rural Centers
—	20.8	13.3	—	7.5	—
—	104.4	102.5	—	1.9	—
26.1	66.7	58.3	—	1.6	6.8
20.1	32.0	21.0	—	5.5	5.5
11.2	103.8	100.3	—	1.8	1.7
6.1	38.5	33.9	2.2	1.5	0.9
5.2	21.7	19.4	1.4	0.2	0.7
3.4	86.2	83.9	—	1.2	1.1
72.1	474.1	432.6	3.6	21.2	16.7
7.7	100.0	91.2	0.8	4.5	3.5

Report of Projects Active as of February 1969, 23 April 1969.

Neighborhood Youth Corps-Out-of-School component, and

TABLE 3
POSITIONS AUTHORIZED FOR DOL URBAN CE
PROGRAMS, BY PROGRAM AND TYPE OF PRO

Type of Program Area	MDTA Institutional Component		Contra
	Number	Percent	Number
JOBS cities:			
In SMSAs of 500,000 or more population in 1960	72,740	91.4	71,9
In SMSAs of under 500,000 population in 1960	675	0.9	4
Total	73,415	92.3	72,4
Other DOL urban centers	6,096	7.7	—
Total, all DOL urban centers	79,511	100.0	72,4
Percent of all authorized positions as of February 1969	97.0		

* U.S. Department of Labor, Manpower Administration, Status Report of Projects

1

TABLE 3

R DOL URBAN CENTERS IN FOUR STUDY MANPOWER
AND TYPE OF PROGRAM AREA, AS OF FEBRUARY 1969*

nt	JOBS Contract Component		Neighborhood Youth Corps Out-of-School Component		New Careers	
nt	Number	Percent	Number	Percent	Number	Percent
	71,926	99.3	17,432	86.4	2,582	96.4
	483	0.7	312	1.6	—	—
	72,409	100.0	17,744	88.0	2,582	96.4
	—	—	2,420	12.0	97	3.6
	72,409	100.0	20,164	100.0	2,679	100.0
	100		90.5		96.4	

Report of Projects Active as of February 1969, 23 April 1969.

TABLE

ENROLLMENT IN DOL URBAN CENTERS IN

BY PROGRAM AND TYPE OF PRO

Funding Source and Program	Number of Enrollees				
	Total, All Centers	JOBS Cities		Other DOL Urban Centers	DO Rur Cen
		In SMSAs of 500,000 or More Population in 1960	In SMSAs of Under 500,000 Population in 1960		
MDTA (Institutional)	4,445	1,916	113	776	1,
JOBS	—	—	—	—	—
NYC (Out-of-School)	2,284	1,993	43	112	
New Careers	4,917	3,377	143	1,305	
					P
MDTA (Institutional)	25,774	21,834	336	2,345	1,
JOBS	15,963	15,782	181	—	—
NYC (Out-of-School)	21,691	16,652	503	2,346	2,
New Careers	1,507	1,411	—	17	
MDTA (Institutional)	30,219	23,750	449	3,121	2,
JOBS	15,963	15,782	181	—	—
NYC (Out-of-School)	23,975	18,645	546	2,458	2,
New Careers	6,424	4,788	143	1,322	1

TABLE 4
NTERS IN FOUR STUDY MANPOWER PROGRAMS,
OF PROGRAM AREA, AS OF FEBRUARY 1969

	Percent Distribution of Urban Enrollees				
		JOBS Cities			
DOL Rural Centers	Total, All Centers	In SMSAs of 500,000 or More Population in 1960	In SMSAs of Under 500,000 Population in 1960	Other DOL Urban Centers	Percent of Current Enrollmen in Urban Centers
CEP Funding					
1,640	100.0	68.3	4.0	27.7	63.1
—	—	—	—	—	—
136	100.0	92.8	2.0	5.2	94.0
92	100.0	70.0	3.0	27.0	98.1
Program Funding					
1,259	100.0	89.0	1.4	9.6	95.1
—	100.0	98.9	1.1	—	100.
2,190	100.0	85.4	2.6	12.0	89.9
79	100.0	98.8	—	1.2	94.8
Total					
2,899	100.0	86.9	1.7	11.4	90.4
—	100.0	98.9	1.1	—	100.
2,326	100.0	86.1	2.5	11.4	90.3
171	100.0	76.6	2.3	21.1	97.3

III. DEVELOPMENT OF SAMPLE DESIGN

INTRODUCTION

3.1 The study objectives pose conflicting problems for sample design.[1] As reflected in the preliminary OEO-DOL specifications, major emphasis is placed on sufficient sampling and auxiliary data collection per area to permit area-by-area analysis. The projection of study findings to a national level is desirable but lower in priority. The design approach followed is to choose the method of sampling areas, but not their number, with a view to national projections. Speculated levels of sampling error for each of the two types of analyses with the preliminary specifications were then developed as a basis for study recommendations.

DESIGN APPROACH

Study Objectives for Design

3.2 For developing the sample design, the major study objectives are considered to be to:

a. Estimate the benefits and costs of the study programs

b. Identify preferred programs and/or combinations of components for various types of individuals served

c. Analyze the contribution of individual components of programs to benefits and cost

d. Identify the barriers to program effectiveness and how they may best be treated.

---

[1] See Appendix B for some bibliographic notes.

These objectives, of course, subsume a number of questions for analysis. For example:

1. What program components and/or programs seem to be most effective for target individuals having specified characteristics?
2. What are the impacts of area-environment constraints (e.g., hiring practices, location of jobs in relation to residences of target individuals) on program effectiveness, and what program approaches seem most effective under given constraints?
3. How can the allocation of program funds be improved?

The study is also expected to provide additional insight, based on longitudinal observation, as to the ways in which program, area, and individual characteristics interact to yield observed results. To provide a basis for designing the study sample, these descriptions of the study objectives must be translated into specific estimates and/or analyses to be produced in the study.

Study Estimates for Design

3.3 A wide variety of analyses will be carried out in the study. To help put these in perspective, it is useful to consider what might be an idealized set of analyses. From a programmatic point of view, decisions that might be made with the help of the study analyses implicitly involve a prediction as to the impact that the program changes will have on the target populations involved. From this point of view, therefore, one idealized set of study analyses would be to permit program planners and administrators to estimate the benefit and cost implications of possible program changes.[2/] Such changes may be broadly classified by type of change as follows.

Type of Program Change	Relevant Study Objective
1. Change in level of program funding, no other change	a
2. Change in composition of program enrollee population/target population	b
3. Redesign of program/components, other than change in funding or target population.	c,d

---

[2/] This corresponds to the definition of program effect in an analysis of variance model if the absence of interaction between factors is not assumed. That is, a weighted average over all classes, the weights being the sizes of the classes in the target population.

To the extent to which decisions may be taken with regard to an individual program, only estimates for that program are needed; for shifts between programs, estimates of the differences between programs are needed. Thus, a series of estimates might be envisioned for each program showing benefits and costs for its target population and subgroups of that population, and for that of possible alternative programs. Such estimates of impact would be of interest, also, for components of the programs. Finally, it would be of interest to develop an estimate of the overall benefits and costs that might follow with an "optimum" program mix; i.e., if the program and program components found to be best in the study for different population groups were combined in a single hypothetical program.

3.4 The sizes and composition of the program target populations by area are not known sufficiently well to serve as the basis for such estimates. [3/] A useful approximation, depending upon the bias of the programs in attracting and/or selecting participants from other target populations, is to make these estimates with references to the population consisting of current program participants. The alternative of computing measures of benefits and costs just for the sample of program participants and controls observed in the study should be carried out, but represents too narrow an analysis for program planning and evaluation—and, consequently, for a design objective. That is , one is interested not only in averages per person, but also in how many persons may be involved.

Implications of Study Objectives for Design

3.5 The fact that measures of program impact are a major interest in the study, and that it would be desirable to project these to a national setting, if feasible,has several implications for the sample design. First, the study areas should not be a selection of special case studies as might be justified in an exploratory study of program methodology. They should, however, reflect a range of area environment and program techniques. This point of view was reflected in a set of criteria suggested by DOL for use in selecting the study areas. Second, the use of probability sampling techniques at each level of sampling rather than judgmental or uncontrolled selection seems desirable. This approach provides conditions necessary for valid application of the tools of statistical inference and hypothesis testing. It is especially attractive since preliminary speculation suggests that national projections can be made from the study within limits of sampling error useful for benefit-cost analysis. Perhaps more important, however, it helps to avoid criticisms of bias, conscious or otherwise, in the selection of study areas, projects and individuals. This can be pointed up by the reference to the basic research design of the study. Since enrollees are to be interviewed on enrollment, the fact that they are the subjects of evaluation will be known to program staffs. This could

---

[3/] The sample of the program target populations to obtain the control group for the study will provide estimates of the size and composition of the target populations, but these estimates are not likely to be sufficiently reliable by area.

be reflected in special selection and/or treatment of enrollees. Some explicit check on this may be possible from data on prior experience of the program, and on the target population. Regardless of this, if favorable outcomes for enrollees are observed in a probability sample of areas, the inference is that what was accomplished in the study areas can be accomplished elsewhere. On the other hand, if the study areas were chosen subjectively, it would be difficult to disprove the argument that this inference should not be drawn. The use of probability sampling is not a panacea for the complex problems of interpreting the data to come from the study. Nevertheless, there seems to be little ground for introducing additional uncertainties by the use of subjective methods of selection.

Estimation Approaches

3.6 Detailed estimates of the type directed to study objectives (c) and (d), and to some extent objective (b), are most efficiently derived from analysis approaches such as the use of regression equations which build in a model structure with corresponding reductions in sample size requirements compared with an approach of making independent estimates for each analysis group of interest. For example, estimates of average post-program earnings per enrollee for each of four race-sex groups can be derived from a regression equation with dummy variates for race and sex with a sample of enrollees only three-fourths as large as would be required to make independent estimates for each group with the same sampling errors.[4] Greater efficiencies are achieved when more factors are involved.

3.7 Such approaches depend on the applicability of the model assumed.[5] Thus, the result just cited for estimates by race-sex group depends on the assumption that program impacts by race and by sex are additive; i.e., there is no interaction between the two factors. If such interaction exists, no improvement in statistical efficiency is achieved by the regression approach. By definition, no interaction between race and sex would mean that the differences in average earnings between males and females would be the same for each of the two race groups, and between races would be the same for each of the two sex groups. The available evidence indicated that this was not likely to be the case, and that interaction would exist.

3.8 A number of earlier evaluation studies were reviewed to obtain some evidence as to the efficiencies which might be achieved by the use of complex estimation techniques within race-sex-age groups. Because of the lack of detailed data in the study reports, it was only possible to reach a general conclusion that regression equations based on as many as 30 variables including race, sex, and age would not be likely to reduce the variance of a variable such as post-program earnings by as much as a half. If applied to make estimates by race, sex, and age group, the gain for each such estimate would presumably be much smaller.

---

[4] See Appendix C.

[5] Certain analyses planned for the study will use cluster analysis techniques which are non-parametric. A typical analysis of this type would be to identify characteristics of enrollees who appear to benefit most from a given program or program component.

Design Approach

3.9 In the light of these considerations, it was concluded that the study sample should be designed to provide independent estimates of program impact for the race-sex-age groups outlined. More detailed analyses—aimed at assessing the impact and psychological characteristics of enrollees, and the demographic, socio-economic constraints—would be based on techniques such as regression and cluster analysis. The depth of analysis feasible would then depend on the structural relationships found to exist.

3.10 Accordingly, the investigation focused on the problem of estimating the benefits of each program for its enrollee population and differences in benefits between programs for enrollees representing matched target populations. Defining benefits for this purpose as post-program changes in annual earnings, it is sufficient to consider annual earnings itself. The types of estimates by program considered are illustrated in Table 5, and program comparisons in Table 6.

DESIGN INVESTIGATION

Approach

3.11 The major objective of the investigation at this point was to establish a reasonably efficient sampling of areas and rough estimates of expected sampling errors. In a two-stage design, the sampling at both stages enters simultaneously into the optimization of the design. On the other hand, it was not possible when the analysis was carried out to be very specific as to a within-area sample design that would be operationally feasible; nor were data available to develop estimates of the variance parameters associated with design features that might be considered. In the absence of such knowledge, an optimum design can only be approximated on the basis of some general considerations. To obtain some insight as to an optimum design, a general line of argument based on a simplified design-concept was therefore explored. The results were then reviewed with regard to the potential impact of different assumptions on the conclusions as to the sampling of areas and the speculated levels of sampling error.

3.12 Since the critical questions concerning the sampling of areas are those of the number of areas to be selected and the probabilities of selection, these were considered first. The motivation for giving some areas a higher chance of selection than others in a study such as this is that they contribute disproportionately to statistics of interest, or have such special significance that one wishes to be certain to include them for other reasons. A number of possible measures come to mind as being suitable for use as a basis for arriving at area selection probabilities, but on which no information is available. One such measure on which information is available is size of area, i.e., number of training opportunities authorized. This is the measure used in the sample design investigation.

ILLUSTRATION OF PROGRAM

Age Group and Type of Enrollee	Estima					
	Job Corps			JOBS		
	Program	Control	Change	Program	Control	Change
All ages						
All enrollees				X*	X	X
White				X	X	X
Non-white				X	X	X
Male				X	X	X
Female				X	X	X
White						
Male				X	X	X
Female				X	X	X
Non-white						
Male				X	X	X
Female				X	X	X

*X indicates an estimate.

TABLE 5
PROGRAM ESTIMATES FOR SAMPLE DESIGN ANALYSIS

Estimated Post-Program Annual Income									
	MDTA(Inst)			NYC(Out-of-School)			New Careers		
Change	Program	Control	Change	Program	Control	Change	Program	Control	Change
X	X	X	X						
X	X	X	X						
X	X	X	X						
X	X	X	X						
X	X	X	X						
X	X	X	X						
X	X	X	X						
X	X	X	X						
X	X	X	X						

Age Group and Type of Enrollee	Est					
	Job Corps			JOBS		
	Program	Control	Change	Program	Control	C
Under 22						
All enrollees	X[**]	X	X	X	X	
White	X	X	X	X	X	
Non-white	X	X	X	X	X	
Male	X	X	X	X	X	
Female	X	X	X	X	X	
White						
Male	X	X	X	X	X	
Female	X	X	X	X	X	
Non-white						
Male	X	X	X	X	X	
Female	X	X	X	X	X	

*Up to 10 percent of participants may be under 22, but this is not a large program group.

**X indicates an estimate.

TABLE 5 (Cont)

timated Post-Program Annual Income									
e	MDTA(Inst)			NYC(Out-of-School)			New Careers*		
Change	Program	Control	Change	Program	Control	Change	Program	Control	Change
X	X	X	X	X	X	X			
X	X	X	X	X	X	X			
X	X	X	X	X	X	X			
X	X	X	X	X	X	X			
X	X	X	X	X	X	X			
X	X	X	X	X	X	X			
X	X	X	X	X	X	X			
X	X	X	X	X	X	X			
X	X	X	X	X	X	X			

TABLE 5

Age Group and Type of Enrollee	Estim						
	Job Corps			JOBS			
	Program	Control	Change	Program	Control	Change	P
22 and over							
All enrollees				X*	X	X	
White				X	X	X	
Non-white				X	X	X	
Male				X	X	X	
Female				X	X	X	
White							
Male				X	X	X	
Female				X	X	X	
Non-white							
Male				X	X	X	
Female				X	X	X	

*X indicates an estimate.

TABLE 5 (Cont)

Estimated Post-Program Annual Income									
	MDTA(Inst)			NYC(Out-of-School)			New Careers		
hange	Program	Control	Change	Program	Control	Change	Program	Control	Change
X	X	X	X				X	X	X
X	X	X	X				X	X	X
X	X	X	X				X	X	X
X	X	X	X				X	X	X
X	X	X	X				X	X	X
X	X	X	X				X	X	X
X	X	X	X				X	X	X
X	X	X	X				X	X	X
X	X	X	X				X	X	X

TABLE 6
ILLUSTRATION OF PROGRAM COMPARISONS FOR SAMPLE DESIGN ANALYSIS

Age Group of Enrollees	Programs Compared
All ages	JOBS, MDTA (Inst)
Under 22	Job Corps, JOBS, MDTA (Inst), NYC (Out-of-School)
22 and over	JOBS, MDTA (Inst), New Careers

3.13 Consider a sample design in which a sample of areas is sampled with the probability of selection allowed to vary by area, and a sample of enrollees within area. For arriving at an optimum design, the total cost available for the study is assumed to be fixed, rather than the preliminary OEO-DOL specifications. The criterion of an optimum design is minimum sampling error for a fixed total cost. The optimization of the sample design then consists in determining values for the numbers of areas, selection probabilities, and numbers of enrollees to be sampled such that the sampling variance of some estimate (estimates) of interest is minimized subject to a given fixed total cost. The results of this optimization, if in a general form, provide guidance as to the number of areas to be used and their selection probabilities determined so as to improve the efficiency of the sample design.

Optimization Analysis

3.14 The sampling model used, applicable to each program, assumes that:

a. There is a fixed set of areas. The probability of selection for areas may be varied by area as appropriate. For simplicity it is assumed that the areas are sampled independently with replacement.

b. Within each of the selected areas, a sample of enrollees is selected from the program. It is assumed that the sample of enrollees is the equivalent of a simple random sample from a large population of all such individuals. The sample sizes within area are assumed to be fixed. The program outcome observed for an individual is assumed to be a fixed number.

3.15 The statistic considered for the analysis is a national projection of a ratio such as average annual post-program earnings per enrollee or the benefit-cost ratio for a given program. The ratio is constructed by taking the ratio of a weighted sum of the area by area projections of the numerator variable and a corresponding average for the denominator variable. The weights reflect the selection probabilities. The projection factors may be arbitrary. Specifically, denote the ratio by r, then

$$r = x'/y' \tag{3.1}$$

where

$$x' = \sum_{h=1}^{q} \frac{1}{P_h} \left(\frac{N_h}{n_h} x_h\right) \tag{3.2}$$

and

$$y' = \sum_{h=1}^{q} \frac{1}{P_h} \left(\frac{N_h}{n_h} y_h\right) \tag{3.3}$$

In these expressions, x and y denote the numerator and denominator variables, respectively; q is the number of areas selected; P_h is the probability of selection of area h on a single draw; x_h and y_h are the totals of x and y, respectively, for the sample of enrollees in area h; n_h is the number of enrollees in the sample for area h, and the ratio (N_h/n_h) is the projection factor for the area. The projection for x, say, in area h may be expressed as

$$\frac{N_h}{n_h} x_h = N_h \bar{x}_h \qquad (3.4)$$

where $\bar{x}_h$ is the mean per enrollee. If N_h represents the size of an assumed population to which the sample findings are being projected, the right-hand side of (3.4) expresses the customary method of projection used by analysts.

3.16 The costs considered for the optimization are those affected by the sample design. It is assumed that these costs can be approximated by the simple linear cost function

$$C = c_1 q + c_2 q \sum P_h n_h \qquad (3.5)$$

where c_1 is a unit overhead per area and c_2 is a unit cost per enrollee in the sample. The cost c_1 would cover elements of the study such as LS&R costs for team visits to an area to arrange for the study and the collection of data on the area, and NORC supervisory costs. The cost c_2 would cover elements such as sampling, interviewing, and data processing per enrollee. In both cases, the costs should be interpreted as marginal or incremental costs for an added unit in the sample (area or enrollee).

3.17 If the sample is made self-weighting, it follows from known theory that the selection probabilities which minimize the variance of the estimator (3.1) subject to the fixed cost are given by

$$P_h = \frac{\sqrt{A_h}}{\sum_{h=1}^{Q} \sqrt{A_h}} \qquad (3.6)$$

where, approximately,

$$A_h \doteq N_h^a \delta_h \tag{3.7}$$

and δ_h is a measure of homogeneity of enrollees within an area, similar to an intraclass correlation, with regard to the variable of interest (e.g., post-program earnings).[6] Details are given in Appendix D. If the δ_h are approximately constant over the range of areas, as might not be unreasonable in this case, the optimum probabilities for the sampling of areas are proportional to the N_h. Thus, for projections to the level of annual program operations, the optimum probabilities would be proportional to the size of program in terms of number of enrollees. If the size of program is correlated with the size of the program target population over the range of areas, this would also be the case for projections to the target population. Alternative possibilities for the behavior of the δ_h over the range of areas can be speculated. For example if, as the size of program increases the enrollees become more and more heterogeneous, then the δ_h would decrease with increasing size of program. If the δ_h decrease inversely as the N_h increase, so that the product $N_h \delta_h$ is approximately constant over the range of areas, the optimum selection probabilities would be proportional to the square root of program size $\sqrt{N_h}$. Equal probabilities of selection would be optimum if the δ_h decrease inversely as the square of N_h, so that the product $N_h \delta_h$ would decrease with increasing N_h. In a wide variety of applications in sampling human populations, it has been found empirically that the rate of decrease in δ is usually small enough so that the product $N\delta$ increase with N. Thus, the optimum selection probabilities would be somewhere between proportional to N_h and proportional to $\sqrt{N_h}$.

Considerations for More Than One Program

3.18 The optimum probabilities for selecting areas following this approach vary for the different programs, according to the distribution of the different programs among the areas. The approach can be readily extended to the case of estimating inter-program differences, but the results are not informative since there still is the problem that the optimum selection probability for an area depends on the pair of programs compared. This was of particular concern because of the variation in the scale of the different programs. If the ratios for different programs over the set of areas are not highly correlated, the variance of an inter-program comparison is the sum of the variance of the ratio for each of the two programs compared. Thus, one may as well deal directly with the individual program variances themselves. The approach taken was as follows. Consider replacing the optimum probability for an area in a given program, say P_h, by

---

[6] M.H. Hansen and W.N. Hurwitz, "On the Determination of Optimum Probabilities in Sampling," Ann. Math Statist., XX (1949), 426-432.

$$P'_h = (1 + f_h) P_h$$

where f is some adjustment factor which can be positive, negative, or zero if there is no change. With the assumed sampling model, the effect on the contribution to the total variance of the program ratio arising from the sampling of areas is to multiply it by the factor

$$\sum_{h=1}^{Q} \frac{1}{1+f_h} .$$

An empirical investigation was then conducted to examine the effect of some approaches commonly suggested for the related problem of sample allocation when there is more than one statistic of interest.[7] For this purpose, the universe of study areas proposed in Section II was used. Probabilities of selection for each of the five study programs were computed on the two rules: proportional to size and proportional to the square root of size. These were then changed under different rules such as using the largest value of the individual probabilities, the average value, etc. Generally speaking, there was a high correlation between the probabilities assigned under the different rules. The effects of different rules on the selection probabilities for the individual areas were also considered. All of the rules tended to improve the selection probability of areas in which any program was large. This had the useful side effect of increasing the representation in the sample of areas of interest to the Washington program staffs (see Section IV).

Summary of Investigation of Sampling the Study Areas

3.19 From the investigation, it was concluded that the following compromise is likely to give reasonable results.

---

[7] A mathematical programming approach was not followed because of the interest in an individual area assessment. An assessment of the implications of the compromises for design efficiency is given in Appendix D.

a. For each of the programs, separately, compute the probability of selecting the individual areas. This probability is taken to be the proportion of the total program size over all areas that is accounted for by the given area.

b. Assign to each area the maximum value of the individual program probabilities.

c. Normalize the resulting set of probabilities to add to unity.

d. Exclude the New Careers program from the computation of selection probabilities.

The final choice of a procedure was based on the results for the largest areas, which it was felt should come into the sample with certainty. The New Careers program is excluded because it is the smallest of the programs and appears in only a limited number of the areas in the proposed study universe. Thus, to include it affects the other program estimates without really helping those for New Careers. If estimates for the New Careers program need to be made in the same detail and depth of analysis as for the other programs, consideration should be given to adding a few additional study areas for this purpose.

3.20 A question may be raised as to why the design approach followed was adopted rather than to stratify the areas by size and to explore optimum allocation of the sample of areas among the strata. The answer is that with as few as 10 areas to be selected there would only be a limited amount of stratification carried out and it seemed inefficient to waste it on size of area when, from a theoretical point of view, the same effect could be achieved by varying the selection probabilities.[8/]

Recommended Design for Sample of Areas

3.21 On the basis of the analysis carried out along the lines just indicated, the design approach recommended is as follows:

a. Assign selection probabilities to each area as described in paragraph 3.19.

b. Stratify the areas on the basis of background considerations. Establish as many strata as there are areas to be selected. Equalize the strata in terms of the sum of the selection probabilities for the areas which constitute them.

c. Select one area from each stratum in accordance with the assigned probabilities.

[8/] Extensive use of "deep" stratification or experimental design techniques were not considered because of the small number of areas to be selected, and the uneven experience shown by research on them. See, for example, E.C. Bryant, H.O. Hartley, and R.J. Jessen, "Design and Estimation in Two-Way Stratification," J. Amer. Statist. Assoc., 55(1960), 105-124.

d. Use as measures of program size authorized positions by program as of February 1969. The assumption in this is that there is expected to be a fairly high year-to-year correlation in program size by area. If it is known what changes will be made in FY70, compared to FY69, some adjustment should be made in the assigned probabilities.

The recommendation to stratify to the point where only one area is selected per stratum is based on the judgment that with as few as 10 sample areas, priority in the design of the sample of areas should be given to improving national projections of the study data through more stratification rather than to unbiased estimates of between-area variances. Thus, with 10 sample areas, two of which are certainty areas, a total of 8 strata of non-certainty areas could be set up assuming that deep stratification is not used. If there were to be at least two areas per stratum, this would allow a maximum of four strata. This judgment is reinforced by the proposed approach of building up the analysis of the total sample from an area-by-area analysis. The recommendation to equalize the size of strata is based on theoretical analysis combined with empirical experience which shows this to be efficient. These recommendations relate to the type of sampling unit represented by the study areas; i.e., a cluster of individuals/programs which are the elementary unit of analysis for the study. When the sampling unit is the elementary unit of analysis, other recommendations might be appropriate.

3.22 A 10-area design following these recommendations was carried out and a sample of 10 areas selected. This is described in Section IV.

SPECULATED SAMPLING ERRORS

3.23 In the discussion below, the approach followed to develop speculated sampling errors with the preliminary specification for 10 areas and the numerical estimates obtained are taken up first. This is followed by a discussion of the application of the speculated sampling errors to the study design.

Approach

3.24 The data most directly relevant to developing numerical estimates of variance parameters for assessing the preliminary specifications for 10 areas and 10,000 individuals in the study are previous data on annual post-program earnings of enrollees on an individual person basis, and comparable data for non-participants who might serve as control cases or, failing this, data on pre-program earnings of controls. Data approximating these specifications were made available by OEO from a national evaluation study of Jobs Corps terminees and by DOL from a national study of NYC (O/S) terminees. Four generalized curves were developed by analysis of these data, as described in Appendix E:

a. A curve showing the relvariance (square of the coefficient of variation) between program enrollees within project, for annual post-program earnings; as a function of annual post-program earnings; and a corresponding curve to represent a comparable relvariance for control cases (Figure E.1)

b. A curve showing the total relvariance of annual post-program earnings per enrollee as a function of the within project relvariance; and a corresponding relationship for control cases (Figure E.2).

The speculated levels of sampling error that might be expected in estimates from these four generalized curves are described in Appendix D. The qualifications on these speculated sampling errors are discussed in Appendices D and E.

The Speculated Sampling Errors

3.25 Two types of sampling errors were speculated. The first type appears in tables and curves labeled "Fixed Set of Areas." These sampling error figures do not include a between-area component of variance and are, therefore, conditional sampling errors. They are applicable for tables and analyses based on the total sample in all areas, but in which the 10 study areas are treated as a universe. The second type appears in tables and curves labeled "National Projection." These sampling error figures do include a between-area component of variance for the sampling of areas. They are applicable for tables and analyses in which the data from the 10 study areas are expanded to national estimates (on the basis of their weights), and the analysis carried out on these national estimates. Sampling errors shown for a given level of earnings on a fixed set of areas basis decrease inversely as the square root of the sample size. Those for national projections decrease more slowly, because of the between-area component of variance, and cannot decrease below the sampling error due to areas. For estimates of average annual earnings for either program terminees or controls it is assumed that two of the 10 areas are self-representing so that the between-area component of variance arises from the sampling of eight of the ten areas. In addition, for program-control and program-program comparisons it is assumed that earnings have a 0.3 correlation at the area level so that the between-area component of variance for these comparisons is reduced by the factor

$$(1 - \rho) = (1 - 0.3) = 0.7$$

where ρ is the correlation. Details are given in Appendix D.

3.26 Because of the nature of the data available for the sampling error analysis, it was not possible to separate out of the variance between projects the components of variance between areas and between project within area. Therefore, in the speculated sampling errors, the entire between project variance is treated as the between-area variance. To the extent to which projects are sampled within area, the "Fixed Set of Areas" sampling errors are under-estimates of the sampling errors to be expected. Regardless of this, the "National Projection" sampling errors are overstatements of the sampling errors to be expected whenever more than one project is in the sample from an area.

3.27 Despite the qualifications and uncertainties that are attached to the speculated sampling errors, it is felt that they provide a reasonable indication of the level of sampling errors to be expected in independent estimates from the study of average post-program earnings of enrollees and the corresponding earnings of control cases. (See Table E.7, Appendix E.)

3.28 Sampling Errors for Estimates of Average Annual Earnings. Table 7 shows the speculated sampling errors for independent study estimates at average annual post-program earnings per enrollee in a given program, according to the level of the average earnings and the number of terminees on which the average is based. Sampling errors are shown for both the "Fixed Set of Areas" and "National Projections". Table 8 shows the speculated sampling errors for comparable estimates for the program control group. The figures in this and succeeding tables are shown to the last digit for convenience in interpolating for other values--obviously they are not precise to the last digit.

3.29 The estimates in Tables 7 and 8 are applicable to averages for a race-sex-age group, where age is either under 22 or 22 and over, and can be used for estimates for combinations of such groups. Estimates are shown by level of average earnings rather than by race-sex-age because, in the experience analyzed, variability in earnings between individuals in a group, from a sampling point of view, seemed to be characterized more closely by level of average annual earnings than by race-sex-age. Thus, a group of non-white male enrollees whose annual post-program earnings averaged, say, $4,000 a year would exhibit variability in earnings between individuals more like that of, say, a group of white male enrollees with the same average annual post-program earnings than a group of non-white male enrollees whose post-program earnings averaged, say $2,000. It may be noted that the sampling errors with a given sample size increase with increasing average earnings but that the coefficients of variation (sampling error divided by average earnings) decrease. This latter is due to the fact that as the average earnings level of a group increases, the group tends to exhibit less variability in work patterns between the individuals.

3.30 The samples sizes of 1,000, 500, and 250 terminees for which sampling errors are shown may be considered as roughly corresponding to the number of reports expected in the full sample (all areas) for a given program, a half-sample,

TABLE 7

SPECULATED SAMPLING ERROR OF AVERAGE ANNUAL POST-PROGRAM EARNINGS PER TERMINEE, FIXED SET OF AREAS AND NATIONAL PROJECTION, BY NUMBER OF TERMINEES IN SAMPLE*

Average Annual Post-Program Earnings per Terminee (Dollars)	Speculated Sampling Error, (Dollars) of Estimated Average Annual Post-Program Earnings per Terminee, Fixed Set of Areas, if the Number of Terminees Is			Speculated Sampling Error (Dollars) of National Projection of Average Annual Post-Program Earnings per Terminee, if the Number of Terminees Is		
	250	500	1,000	250	500	1,000
4,000	222	157	111	257	203	170
3,200	189	133	94	219	174	146
2,600	164	116	82	192	152	128
2,000	137	97	69	163	131	111
1,600	118	83	59	141	114	98
1,200	95	67	47	116	94	81
1,000	82	58	41	99	83	72
800	69	49	34	85	70	61
600	55	39	28	69	56	49
500	49	35	24	60	49	43

* Figures for number of terminees are for individuals from whom data are obtained, not initial samples. See text for description of sample of areas.

TABLE 8

SPECULATED SAMPLING ERROR OF AVERAGE ANNUAL EARNINGS PER CONTROL CASE, FIXED SET OF AREAS AND NATIONAL PROJECTION, BY NUMBER OF CONTROL CASES IN SAMPLE*

Average Annual Earnings per Control Case (Dollars)	Speculated Sampling Error (Dollars) of Estimated Average Annual Earnings per Control Case, Fixed Set of Areas, if the Number of Control Cases Is			Speculated Sampling Error (Dollars) of National Projection of Average Annual Earnings per Control Case, if the Number of Control Cases Is		
	250	500	1,000	250	500	1,000
4,000	248	175	124	279	218	179
3,200	227	161	114	253	195	159
1,600	202	142	101	224	173	141
2,000	172	119	84	189	148	121
1,600	143	101	72	163	128	106
1,200	116	82	58	133	105	88
1,000	101	72	50	117	93	78
800	83	59	42	87	78	60
600	68	48	34	79	63	53
500	60	42	30	69	55	46

* Figures for number of control cases are for individuals from whom data are obtained, not initial samples. See text for description of sample of areas.

or a quarter-sample, respectively, if there were 20 percent attrition from the original sample over the life of the study. Illustrative program terminee groups corresponding to these sample sizes are indicated in the following tabulation.

Sample Size		Illustrative Program Group
Number	Fraction of Sample Size	
1,000	1	Age (under 22, or 22 and over)
500	$\frac{1}{2}$	Race-age; sex-age
250	$\frac{1}{4}$	Race-sex-age

This assumes that the under 22 years of age and 22 years of age and older are treated as separate groups for the JOBS and MDTA (Inst.) programs, and that only one age group or the other is present in the other programs.

3.31 Speculated sampling errors for other sizes of sample may be derived from Tables 7 and 8 by interpolation. Figures 1 and 2 present sampling error curves useful for this purpose. These curves show the speculated sampling errors for estimated average annual post-program earnings per terminee as a function of sample size on a fixed set of areas and national basis, respectively, corresponding to four assumed levels of average annual post-program earnings. The question of allocation of the total sample among the programs is discussed further below.

3.32 Program-Control Group Comparisons. Tables 9 and 10 show the speculated sampling errors on a fixed set of areas and national basis, respectively, for independent study estimates of average increase in annual post-program earnings for enrollees in a given program compared with their controls. Each table has three parts, corresponding to the three levels of sample size, as follows.

Part	Sample Size
A	250 terminees and 250 controls
B	500 terminees and 500 controls
C	1,000 terminees and 1,000 controls

Each part is a double-entry table in which the columns correspond to an assumed level of average annual post-program earnings per terminee and the rows to an assumed level of corresponding average annual earnings per control case. The difference between the levels of earnings specified by a given

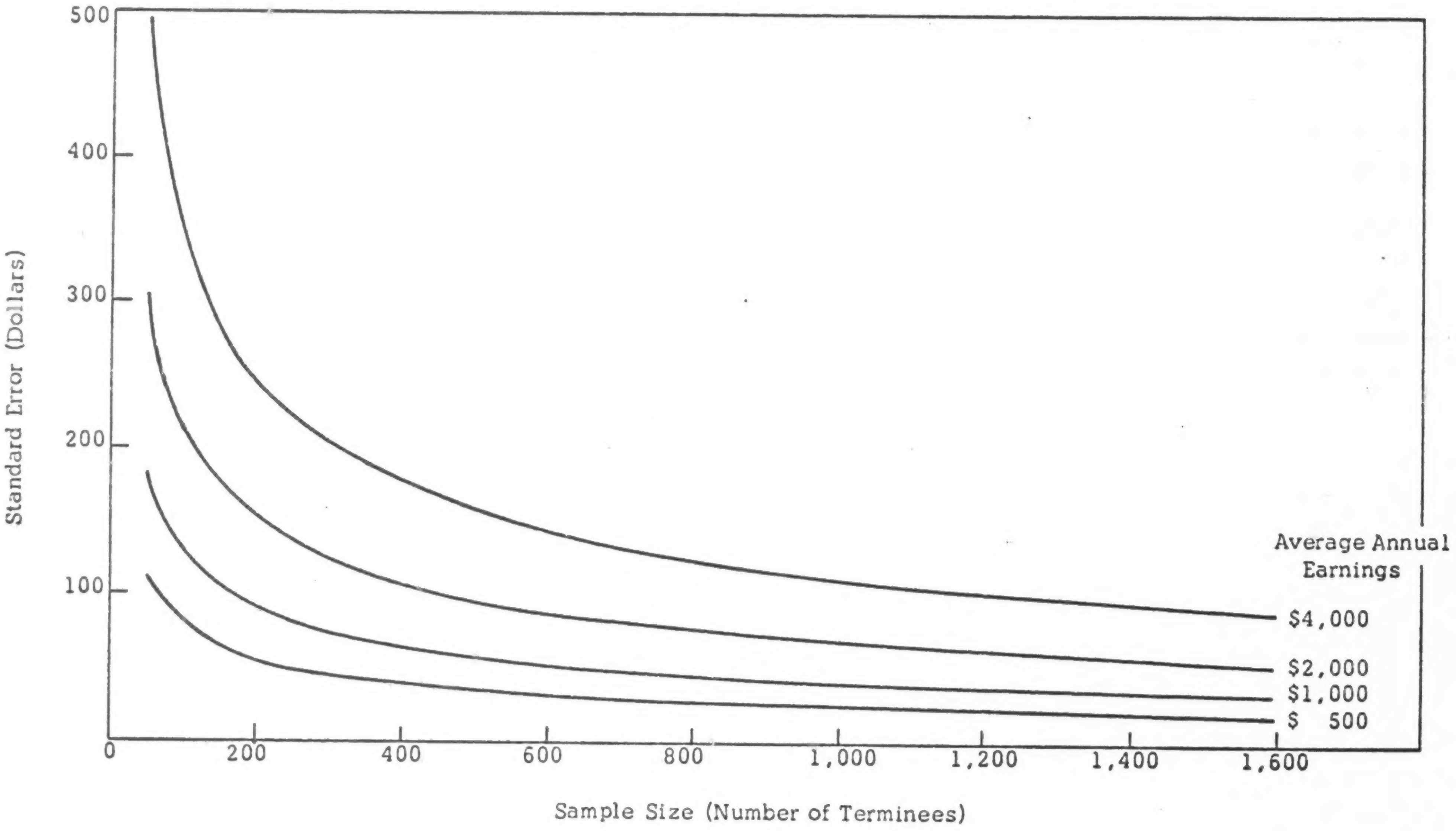

FIGURE 1. SPECULATED STANDARD ERROR OF ESTIMATED AVERAGE ANNUAL POST-PROGRAM EARNINGS PER TERMINEE, FIXED SET OF AREAS, BY LEVEL OF EARNINGS

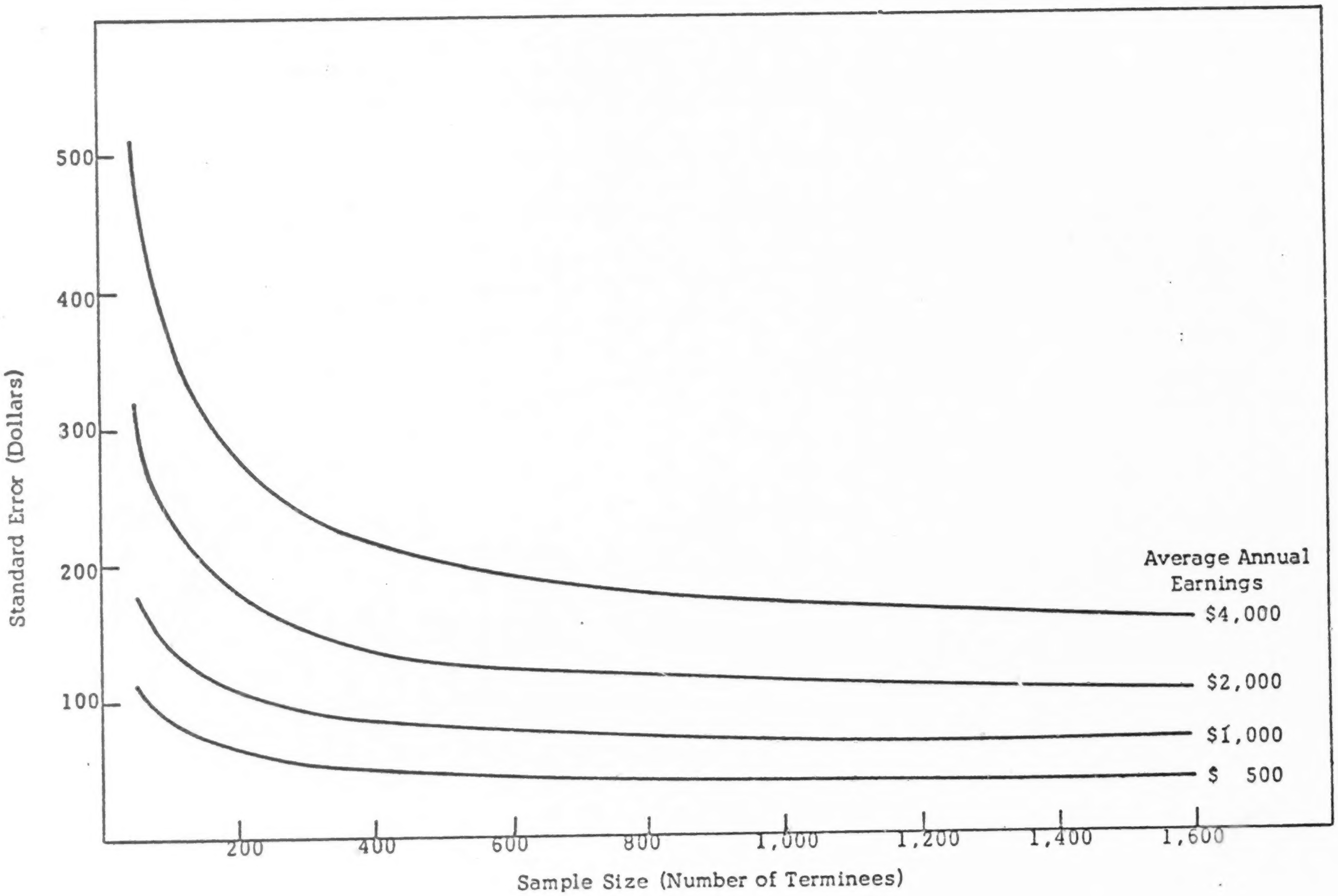

FIGURE 2. SPECULATED STANDARD ERROR OF NATIONAL PROJECTION OF AVERAGE ANNUAL POST-PROGRAM EARNINGS PER TERMINEE, 10 AREA DESIGN, BY LEVEL OF EARNINGS

TABLE 9

SPECULATED SAMPLING ERROR OF ESTIMATED AVERAGE INCREASE IN ANNUAL POST-PROGRAM EARNINGS PER TERMINEE, FIXED SET OF AREAS

a. 250 Terminees and 250 Control Cases

Average Annual Earnings per Control Case (Dollars)	Speculated Sampling Error (Dollars) of Estimated Average Increase in Annual Post-Program Earnings per Trainee, Fixed Set of Areas, if Average Annual Post-Program Earnings per Terminee (Dollars) Is									
	4,000	3,200	2,600	2,000	1,600	1,200	1,000	800	600	500
4,000	333									
3,200	318	295								
2,600	300	276	260							
2,000	278	253	235	217						
1,600	264	237	218	198	185					
1,200	250	221	201	180	165	150				
1,000	244	214	193	171	155	139	129			
800	238	207	185	161	145	127	118	109		
600	232	201	178	153	136	117	106	97	88	
500	230	198	175	150	132	112	102	91	82	77

* Figures for numbers of terminees and control cases are for individuals for whom data are obtained, not initial samples.

TABLE 9 (Cont)

b. 500 Terminees and 500 Control Cases

Average Annual Earnings per Control Case (Dollars)	Speculated Sampling Error (Dollars) of Estimated Average Increase in Annual Post-Program Earnings per Trainee, Fixed Set of Areas, if Average Annual Post-Program Earnings per Terminee (Dollars) Is									
	4,000	3,200	2,600	2,000	1,600	1,200	1,000	800	600	500
4,000	235									
3,200	225	215								
2,600	212	196	184							
2,000	197	179	166	153						
1,600	187	168	154	140	131					
1,200	177	157	142	127	116	106				
1,000	172	151	136	121	110	98	91			
800	168	146	131	114	102	90	83	77		
600	164	142	126	108	96	82	75	68	62	
500	163	140	124	106	93	79	72	64	58	55

* Figures for numbers of terminees and control cases are for individuals for whom data are obtained, not initial samples.

TABLE 9 (Cont)

c. 1,000 Terminees and 1,000 Control Cases

Average Annual Earnings per Control Case (Dollars)	Speculated Sampling Error (Dollars) of Estimated Average Increase in Annual Post-Program Earnings per Trainee, Fixed Set of Areas, if Average Annual Post-Program Earnings per Terminee (Dollars) Is									
	4,000	3,200	2,600	2,000	1,600	1,200	1,000	800	600	500
4,000	166									
3,200	159	148								
2,600	150	138	130							
2,000	139	126	117	108						
1,600	132	118	109	99	93					
1,200	125	111	100	90	82	75				
1,000	122	107	96	85	78	69	64			
800	119	103	92	81	72	63	59	54		
600	116	100	89	77	68	58	53	48	44	
500	115	99	88	75	66	56	51	46	41	39

* Figures for numbers of terminees and control cases are for individuals for whom data are obtained, not initial samples.

TABLE 10

SPECULATED SAMPLING ERROR OF NATIONAL PROJECTION OF AVERAGE INCREASE IN ANNUAL POST-PROGRAM EARNINGS PER TERMINEE*

a. 250 Terminees and 250 Control Cases

Average Annual Earnings per Control Case (Dollars)	Speculated Sampling Error (Dollars) of National Projection of Average Increase in Annual Post-Program Earnings per Terminee if Average Annual Post-Program Earnings per Terminee (Dollars) Is									
	4,000	3,200	2,600	2,000	1,600	1,200	1,000	800	600	500
4,000	366									
3,200	348	323								
2,600	329	303	285							
2,000	307	279	260	240						
1,600	293	263	242	221	207					
1,200	278	246	224	202	186	169				
1,000	271	239	216	192	175	157	148			
800	264	230	206	182	164	144	134	123		
600	258	224	199	173	155	133	122	111	100	
500	256	221	196	169	150	128	117	105	93	88

* Figures for numbers of terminees and control cases are for individuals for whom data are obtained, not initial samples.

TABLE 10 (Cont)

b. 500 Terminees and 500 Control Cases

Average Annual Earnings per Control Case (Dollars)	Speculated Sampling Error (Dollars) of National Projection of Average Increase in Annual Post-Program Earnings per Terminee if Average Annual Post-Program Earnings per Terminee (Dollars) Is									
	4,000	3,200	2,600	2,000	1,600	1,200	1,000	800	600	500
4,000	285									
3,200	266	247								
2,600	252	232	218							
2,000	236	214	199	185						
1,600	225	202	187	171	160					
1,200	215	190	173	157	145	132				
1,000	209	184	167	150	137	123	116			
800	204	178	160	142	128	113	105	97		
600	199	173	154	135	121	105	96	87	78	
500	197	170	152	132	118	101	92	82	73	69

* Figures for numbers of terminees and control cases are for individuals for whom data are obtained, not initial samples

TABLE 10 (Cont)

c. 1,000 Terminees and 1,000 Control Cases

Average Annual Earnings per Control Case (Dollars)	Speculated Sampling Error (Dollars) of National Projection of Average Increase in Annual Post-Program Earnings per Terminee if Average Annual Post-Program Earnings per Terminee (Dollars) Is									
	4,000	3,200	2,600	2,000	1,600	1,200	1,000	800	600	500
4,000	226									
3,200	213	198								
2,600	202	186	175							
2,000	191	173	161	150						
1,600	183	164	152	140	131					
1,200	174	155	141	132	119	108				
1,000	170	150	136	123	113	102	96			
800	166	145	130	117	106	94	88	81		
600	162	140	126	111	100	87	80	73	65	
500	160	139	124	109	98	84	77	69	61	57

* Figures for numbers of terminees and control cases are for individuals for whom data are obtained, not initial samples.

column and row is the assumed increase in post-program earnings compared to the control group. The entry in the corresponding cell of the table is the speculated sampling error of an estimate of the increase in earnings from the study sample. For example, in Part a of Table 9, the cell in the fourth column and sixth row corresponds to an assumption that the true level of average annual post-program earnings for terminees is $2,000 and for their controls is $1,200 per year. The true post-program increase in average annual earnings assumed is thus $800; and the entry in Part a of Table 9 shows that the sampling error in estimating this increase from a sample of 250 terminees and 250 controls in the study is speculated to be $180 (on a fixed set of areas basis).

3.33 Program-Program Comparisons. Estimates of differences in average annual post-program earnings between terminees from two different programs can be expected to have somewhat smaller sampling errors than program-control comparisons based on the same assumptions. Therefore, Tables 9 and 10 may be used as an approximate guide.

Application of Speculated Sampling Errors

3.34 Use of Tables for Enrollee Groups and Programs. The sampling errors presented in the tables indicate the level of sampling error expected for particular enrollee groups and programs under given assumptions as to average outcomes for program terminees and controls. From available studies, it would be expected that average earnings for males would be higher than for females, and for whites would be higher for nonwhites. It is likely that post-program earnings (and controls) over 22 years of age will be higher than for enrollees under 22. It is also likely that post-program earnings for enrollees (and controls) in programs such as JOBS and MDTA (Inst.) will be higher than for NYC (O/S). This will be heavily influenced by the nature of the training occupations. Because the evidence of previous studies as to program outcomes of terminees and performance of controls is mixed, greatest reliance should be placed on the pattern of sampling errors shown. Finally, if economic conditions are good and the labor market favorable, both program terminees and controls can be expected to show higher average earnings than if there is a slow-down or a recession in the economy.

3.35 Interpretation of the Sampling Errors. The sampling error figures shown in Tables 9 and 10 indicate the level of uncertainty expected to be attached to estimates of post-program increases in earnings (see paragraph 3.40). The adequacy of the initial sample specifications for 10 cities and 10,000 individuals is to be judged on the basis of whether the indicated uncertainty is acceptably small. It is suggested that the levels of sampling error indicated by Tables 9 and 10 are analytically useful. This judgment is based on the fact that the indicated levels of uncertainty are smaller than changes in income large enough so that they are of interest, and so that there can be confidence in the estimated changes. Taking up this latter point first, from purely technical considerations there is a limit set by non sampling error (e.g., errors in obtaining income) below which the contribution of possible measurement bias to the measure of uncertainty of estimated income changes cannot be ignored, but must be considered along with sampling variability. For example, suppose that a study were conducted which showed a $50 a year post-program increase in earnings, and that the study sample were sufficiently large so that this increase was statistically significant. The argument here is that there is not sufficient accuracy in the tools for measuring total annual income to feel much confidence in a difference that small. And, moreover, the significance of such a difference for either the individuals in the program target population or for program decision-making is questionable. This suggests that post-program increases of less than a few hundred dollars a year are not likely to be of analytic interest.

3.36 Least Significant Differences. One measure of the adequacy of a sample that has frequently been used is the least significant difference. Thus, for example, suppose that estimates for program increases on the order of twice their sampling errors (or greater) would be considered statistically significant (i.e., greater than zero). The following table shows approximate least significant differences for program-control comparisons, assuming a one-sided test. The figures in this table may be converted to approximate increases in hourly rates, assuming full-time employment, by dividing by 2,000.

TABLE 11

LEAST SIGNIFICANT INCREASES IN PROGRAM EARNINGS

Average Annual Earnings per Control Case	Least Significant Increase in Annual Post-Program Earnings			
	All Terminees *		Race-Sex-Age Group **	
	Fixed Set of Areas	National Projection	Fixed Set of Areas	National Projection
$3,000	$240	$330	$500	$550
$2,000	$180	$210	$380	$430
$1,000	$120	$170	$250	$280
$ 500	$ 70	$110	$140	$170

*Assumes estimates based on 1,000 terminees and 1,000 controls.

**Assumes independent estimates based on 250 terminees and 250 controls in a group.

3.37 OC Curves. A more useful measure of the adequacy of a sample from the point-of-view of significance testing is the probability that if there is a given true post-program increase it will be detected statistically in the study. The function which gives this probability is the OC (Operating Characteristic) curve of the study. The following table shows approximate post-program increases in average annual earnings for all terminees that have specified probability of being detected in the study, on a fixed set of areas and national projection basis, assuming a one-sided test.

TABLE 12
INCREASES IN POST-PROGRAM EARNINGS THAT WILL BE DETECTED WITH SPECIFIED PROBABILITIES

Probability That Post-Program Increase in Earnings Will Be Detected Statistically in Study	Increase in Annual Post-Program Earnings That Will Be Detected with Given Probability if the Average Annual Earnings of Controls Is *		
	$500	$1,000	$2,000
All Terminees--Fixed Set of Areas			
.50	80	110	190
.80	100	170	270
.90	120	200	320
.95	140	220	370
All Terminees -- National Projection			
.50	100	180	250
.80	160	260	390
.90	190	300	470
.95	210	320	530

* Assumes estimates based on 1,000 terminees and 1,000 controls.

A probability of 0.50 corresponds to a 50-50 chance, or even "odds," that the indicated increase in earnings will be detected statistically. A probability of 0.80 corresponds to chances of 4 out of 5, or odds of 4 to 1, that the increase will be detected. A probability of 0.90 corresponds to chances of 9 out of 10, or odds of 9 to 1; and a probability of 0.95 corresponds to chances of 19 out of 20, or odds of 19 to 1.[9,10] Again, it is suggested that the levels of sampling error indicated by Tables 9 and 10 are analytically useful in the sense that the odds are good that post-program increases in earnings likely to be of interest will be detected statistically in the study.

---

[9] The test on which the table is based is calibrated so that if there were no true increase in average annual post-program earnings of terminees compared with controls, the odds would be 19 to 1 against finding a statistically significant difference in the study in favor of the terminees.

[10] It may be noted that the increase in post-program earnings for which there is a 0.80 probability of detection is about 50 percent higher than the least significant differences for the same average annual earnings of controls.

Implications for Benefit-Cost Ratios

3.38 The coefficients of variation (relative sampling errors) for estimated benefit cost ratios can be expected to be somewhat less than those for estimates of increases in post-program earnings. Let

$$r = B'/C'$$

denote a benefit-cost ratio estimated from the study, where B' is an estimate of benefit and C' an estimate of corresponding cost. Then the relvariance (square of the coefficient of variation) of r is, with the usual approximation for a ratio estimate

$$V_r^2 = \frac{\sigma_r^2}{R^2} = V_{B'}^2 + V_{C'}^2 - 2\rho_{B'C'} V_{B'} V_{C'}$$

where σ_r^2 denotes the variance of r, and R the true benefit-cost ratio being estimated. $V_{B'}^2$ and $V_{C'}^2$ denote the relvariances of the estimates B' and C', respectively, and $\rho_{B'C'}$ the correlation between the estimates. Data were not available for deriving estimates of the magnitudes of $V_{C'}^2$ and $\rho_{B'C'}$ which might be encountered in the study for different programs and population groups. To obtain some guidance as to the levels of sampling error which might be found for estimated benefit-cost ratios, suppose that the benefit-cost ratios for enrollees are not correlated with their program costs. This is a relatively unfavorable assumption, although it has been observed in other studies for enrollees, particularly at the upper end of the cost distribution within a program. With this assumption, V_r^2 may be approximated as

$$V_r^2 \doteq V_{B'}^2 - V_{C'}^2 = V_{B'}^2 \left(1 - \frac{V_{C'}^2}{V_{B'}^2}\right) = V_{B'}^2 \left(1 - \rho_{B'C'}^2\right)$$

where $\rho_{B'C'}^2 = V_{C'}^2 / V_{B'}^2$.

$V_{B'}^2$ will be somewhat less for a population group that $V_{E'}^2$, the relvariance for post-program changes in average annual earnings given in Tables 9 and 10, since in deriving the latter no benefits were counted for terminees who went back to school or entered into military service and no benefits other than earnings were included, while there is reason to expect $V_{C'}^2$ to be substantially smaller than $V_{E'}^2$ (see Appendix D). Under these conditions, the relative sampling errors of estimated benefit-cost ratios for particular programs and population groups might be expected to be on the order of say, 80 percent, of those indicated in Tables 9 and 10 for the corresponding change in average annual post-program earnings.

3.39 Speculating on the basis of these assumptions, it appears that the relative sampling errors of estimated benefit-cost ratios might be as high as 10 to 20 percent on a fixed area basis, and 15 to 25 percent for national projections, under some combinations of benefits and costs. The exact relationship between these, and the corresponding impact on the sampling errors to be expected, is somewhat complex and most readily dealt with by considering specific examples.

3.40 Although the speculated sampling errors appear to be high compared to the specifications for precision ordinarily met in survey studies, it is suggested that they are useful for purposes of the study analysis in view of the uncertainties arising from the assumptions made in defining the benefit-cost ratio itself. For example, a 20 percent sampling error in observed benefit cost ratios would have the implications summarized in the following tabulation.

If the observed benefit-cost ratio is	The chances are about 2 out of 3 that the difference between the observed ratio and the benefit-cost ratio from a study of all program enrollees would be less than
1	0.2
2	0.4
5	1.0
10	2.0

The uncertainty due to sampling which is illustrated by this tabulation is relatively small compared to that arising from other sources of uncertainty which affect the estimated benefit cost ratios. Among such factors which affect the benefit-cost ratio are the assumptions as to the patterns of benefits to be projected for time periods not directly observed, the length of the time horizon over which benefits are projected, and the choice of an appropriate rate for discounting future benefits.

SAMPLE ALLOCATION WITH A FIXED TOTAL COST

3.41 The preceding discussion of the sampling errors which might be expected with the preliminary OEO-DOL specification for 10 areas and 10,000 individuals has been generalized so as to permit the evaluation of alternatives in allocating the 10,000 individuals among areas, programs, and population groups. The question of the allocation of the sample of individuals is discussed in the following paragraphs.

3.42 The sample allocation is considered first without the New Careers program. Then, the sampling questions concerning New Careers are discussed. The approach is that of optimization subject to a fixed total cost.

Study Sample Excluding New Careers Program

3.43 There is not sufficient evidence as to differences from area to area in sampling variances between program enrollees, from program to program within area, or from one race-sex-age group to another within program in a given area, to provide a strong case for departing from equal sample sizes by area, by program within area, and by race-sex group within programs. With regard to differences between programs, it is likely that estimates of post-program earnings for enrollees in the JOBS and MDTA (Inst.) programs will have smaller coefficients of variation but larger sampling errors than for the Job Corps and NYC (O/S) programs. OEO and DOL staff have indicated a preference for using the latter as the criterion, and this would argue for larger samples for the JOBS and MDTA (Inst.) programs. However, it is likely that there will be a higher attrition rate in the sample for the under 22 years of age group than for the 22 and over group, and this would argue for larger initial samples for the Job Corps and NYC (O/S) programs. It was pointed out earlier in this section that of the two age groups:

a. Under 22

b. 22 and over

only the first is essentially applicable to Job Corps and NYC (O/S), and only the second to New Careers. If JOBS and MDTA samples are split into these two age groups, as they must be for cross-program comparisons with the other programs, they will essentially have only a half-sample available for such comparisons. It is unlikely that this loss in sample size can be anywhere near recaptured by analytic techniques applied to a half-sample. Because of the wide variation in the target population over 22, due in part to the breadth of the age range covered, this does not seem to be acceptable.

3.44 Thus, if an initial sample group of size n is used for each of the Job Corps and NYC (O/S) programs, the implications are that

a. Samples of size 2n should be used for JOBS and MDTA (Inst.)

b. An initial sample on the order of 1.4 should be used for a separate control group for each of the two age groups.

Details are given in Appendix D. The specification for the control groups assumes that:

a. There is essentially common eligibility among the programs within age group so that a single control group within the age limitation can serve for all of the corresponding programs

b. On the order of 30 percent of an initial sample of controls would be lost as controls in the study because they enroll in programs during the study, and there would be about the same attrition for the remaining 70 percent as for program cases.

In terms of cost, this would be the equivalent of a total sample of 8n. If the 10,000 cases were divided this way, the initial sample would consist of 1,250 cases for each of Job Corps and NYC (O/S) and each of the two age groups in JOBS and MDTA (Inst.); and 1,750 cases for each of the two age groups in the control samples. OEO and DOL preliminary specifications set an objective for the study to hold "avoidable" attrition in the initial sample to 20 percent over the life of the study. (This excludes "unavoidable" attrition such as that due to the death of sample individuals during the study.) If this objective is achieved, and with the preceding asumptions as to losses of control cases, the final sample would be 1,000 cases for each of the programs and control groups. Because the attrition in the sample will not occur all at the beginning, it may be possible to use statistical estimation procedures to take advantage of the larger samples in the earlier interviews of the study sequence for aggregate estimates of post-program income changes (i.e., estimates for which the full longitudinal data are not required for each individual).

New Careers Program

3.45 If a corresponding sample allocation were included for the New Careers program, the total sample would be the equivalent of 9n. If the 10,000 cases were divided this way, there would be an initial sample of 1,100 for each program group and 1,550 for each control group; and a final sample of approximately 900 cases for each group. If the corresponding sampling errors expected are acceptable, this is one possible alternative. Other alternatives would be to drop the New Careers program from the study, or to increase the study budget (primarily the direct data collection and processing costs) to keep the initial sample for each program group at 1,250. In the latter case, it would be advisable to sample additional areas to strengthen the estimates for New Careers.

3.46 It is suggested that one of these latter two alternatives be adopted. The basis for this suggestion is consideration of what is gained for analysis of the New Careers program by accepting an increase in sampling error for the other programs.

3.47 The distribution of the New Careers program by area is quite different from that of the other four programs (see Section II). Thus, national projections for the New Careers program would be expected to have substantially larger sampling errors than in the case of the other programs. On the other hand, unless additional areas were included in the study order to improve the estimates for New Careers, there would be a loss for the other programs without really satisfying the sampling error objectives for New Careers. Further, because of the small size of the programs, it is unlikely that the required sample could be obtained without either extending the initial interview period substantially or adding other areas to the sample. To add additional areas for New Careers would require a further substantial reduction in the sample of individuals to compensate for the added field costs.

IV. A 10-AREA SAMPLE FOR THE STUDY

INTRODUCTION

4.1 In this section, a stratified 10-area sample for the study is described.[1]

STRATIFICATION VARIABLES

4.2 Three groups of variables were considered for potential use in stratifying the study universe SMSAs for sampling of the ten study areas:

a. Program evaluation data or related current information as to expected program performance

b. Data as to the political and social stresses existing in the areas

c. Socio-economic data for the areas bearing on:

   1. Characteristics of the disadvantaged population

   2. Characteristics of the labor market and the area.

It was intended to make maximum use of the views of OEO and DOL staff in developing the stratification, particularly as reflected in ratings or measures that were not quantified and which therefore could not be effectively incorporated in the statistical techniques for preparing estimates from the study.

---

[1] A 15-stratum design was also explored but is not discussed here.

Program Evaluation Data

4.3 No suitable program evaluation data by SMSA were available for the study to draw on.

4.4 DOL prepared special tabulations for the study which gave dropout rates and employment status at follow-up dates for the FY 1968 MDTA (Inst.) program, and dropout rates for the JOBS (Contract) program since inception. DOL experts indicated that these data were subject to reporting biases, and this was confirmed by detailed review of the tabulations. Accordingly, it was decided that these data would be used only for general checking and review of the strata to be established, and for variance analyses, rather than as stratification variables.

4.5 In addition, OEO and DOL were requested to obtain program ratings and OEO community ratings. The OEO Office of Manpower, Community Action, requested ratings from field staff on three aspects of the areas:

a. CAA quality/image/role vis a vis the poor

b. Attitudes of power structure - business, city hall, unions

c. Quality of program (considering CEP, SES, job development, program linkages, etc.).

Cities were to be rated on each factor as very positive, positive, negative, or very negative. Ratings were obtained for 33 of the 46 cities asked about. In addition, the Office of Evaluation arranged for the study to have access to the ratings of 17 cities on an Employment Impact Index (Any CAA Role) developed by Barss, Reitzel & Associates, Inc., in another study conducted for OEO.

4.6 The DOL Office of Manpower Evaluation obtained ratings of the programs from each of the program staffs. Programs were rated as above average, average, or below average. Ratings were obtained for all cities for the JOBS (Contract) and MDTA (Inst.) programs for 41 cities for the NYC (Out-of-School) program, and for 36 of the 37 cities in which there were New Careers projects. Comparison of these ratings between OEO, DOL and with the JOBS (Contract) and MDTA (Inst.) dropout rates showed substantial disagreements, as might have been anticipated. There clearly are cities on which agreement can be reached as to whether a particular program should be rated as good or bad; and even some cities for which this will be so for all of the programs of interest. However, this was not so for the entire set of cities and programs. The decision made was, again, that these ratings would be used only for general checking and review. It appears that for ratings of the type obtained studies such as the OEO Evaluation Study would be of value to assess the rating processes.

Political and Social Stresses

4.7 In addition to the OEO staff ratings, data on civil disorders in the cities were reviewed as an indicator of political and social stresses. The U.S. Riot Commission (Kerner Commission) Report showed that all the cities of interest had experienced civil disorders in the summer of 1967. More recent data compiled by the U.S. Department of Justice were not retrievable. Thus, no particularly useful discrimination between the areas of interest seemed possible on this basis. Further, it did not appear to be possible to project such distinctions for the future.

Socio-Economic Data

4.8 The preceding review led to basing the stratification on socio-economic data for the areas. The data compiled by area are shown in Table 13.

4.9 There were 42 SMSAs and 10 strata to be established. Following the principle of establishing strata of approximately equal size in terms of the probability measure assigned, two SMSAs were large enough to be selected with certainty. Of the remaining 40 SMSAs, three were large enough so that they would each constitute a stratum in combination with only one or two other areas. This would leave on the order of 34 SMSAs and 5 strata to be established. If these numbers had been larger, the use of some clustering technique would have been explored.[2/] In the present case, it was decided to use more subjective methods. The basic approach was to subjectively review the range of variability on the data items compiled which resulted from trial stratifications. As indicated, three strata were built around larger SMSAs as special combinations. For the remaining five strata, a first classification of areas was established on major industry (manufacturing vs other). This variable was suggested by the interest shown in it by DOL staff. It appeared to lead to a sensible grouping of areas. Further stratification was then based on the variables:

a. Rate of population growth

b. Level of income/hourly wage rates

c. Percent nonwhite population

as seemed appropriate. It should be noted that the set of data items showed considerable (but, of course, not perfect) inter-correlations. No attention was given in the stratification to geographic location or to area size, since these would be incorporated in the selection process. No explicit use was made of the data on poverty populations or crime, the former because of the changes thought to have taken place since 1960 and the latter because of the questions to which the data reporting are thought to be subject. It is to be emphasized that the final stratification is a subjective one, and subject to review and concurrence by OEO and DOL staff.

[2/] See, for example, H.P. Friedman and J. Rubin, "On Some Invariant Criteria for Grouping Data," J. Amer. Stat. Assn. 62 (1967), 1159-1178; G.H. Ball, "Data Analysis in the Social Sciences," Amer. Fed. Information Processing Soc. Conf. Proceedings, Fall Joint Computer Conference, 27 Part 1, 533-560, Spartan Books, Washington, D.C., 1965.

TABLE 13

SOCIO-ECONOMIC DATA FOR STUDY UNIVERSE SMSAs

1. Population (in thousands)
"Population Estimates", SMSA: July 1, 1967 Provisional Estimates
U.S. Department of Commerce, Series P-25, No. 411, December 5, 1968.

2. 1960-65, Percent Change, 1960-65
(SMSA's Metropolitan Area Statistics", Bureau of Census, 1968 Statistical Abstract.

3. 1960 Percent Non-White (2 sets of data)
   a. "Metropolitan Area Statistics" (SMSA) Bureau of Census, 1968 Statistical Abstract
   b. Educational Resources Information Center (Central City) U S. Dept. H.E.W., January 1968, "Profiles of Twenty Major American Cities".

4. 1965 Percent Non-White (Estimated)
Central City Educational Resources Information Center, U.S. Dept. H.E.W., January 1968, "Profiles of Twenty Major American Cities".

5. AFDC Recipients (Percent)
   a. No. of Recipients-1966
   Educational Resources Information Center (Central City)
   U.S. Dept. H.E.W., January 1968, "Profiles of Twenty Major American Cities"
   b. Population-1967 (SMSA)
   "Population Estimates", Series P-25, No. 411, December 5, 1968.

6. Crimes Known to Police, 1967, Rate per 100,000 (SMSA)
"Metropolitan Area Statistics", Bureau of Census, 1968 Statistical Abstract; from Dept. of Justice, Federal Bureau of Investigation "Uniform Crime Reports, 1967".

7. 1960-1963 Dropout Rate (Central City)
(Same as 4)

8.
   a. 1967 Average Annual Unemployment (Percent) for Twenty Major SMSA's and Newark
   b. For others-Unemployment rate as of July 1967.
   (Same as 4)

9. Unemployment Rate (Percent)
Estimated Rate--16-19 years, "Statistical Tables on Manpower, 1968", Reprint of the Statistical Appendix to the 1968 Manpower Report, U.S. Dept. of Labor.

TABLE 13 (Cont)

10.	Unemployment Rate Non-White (Percent) Estimated Rate 1967, (Same as 9)
11.	Hard-Core Unemployed (from Mrs. Regelson, OEO)
12.	1960 Percent of Families in Poverty Areas (Outside Central City) Same as (4).
13.	1960 Percent of Families Below Poverty Level (Outside Central City) Same as (4).
14.	1960 Percent of Families in Poverty Areas (Central City) Same as (4).
15.	1960 Percent of Families Below Poverty Level (Central City) Same as (4).
16.	Funds Allocated, Anti-Poverty (OEO) Programs per person - assumes percent of population below poverty level is same for 67 as 60. Poverty population is arithmetic average of Central City and Outside Central City. Same as (4). (Norfolk data are questionable)
17.	1959 Median Family Income "Metropolitan Area Statistics" (SMSA)
18.	Employment Concentration - Percentage in Two largest industry groups for 1967 "Metropolitan Area Statistics" (SMSA)
19.	1967 Average Weekly Earnings "Employment and Earnings Statistics for States and Areas 1939-67" #1370-5, U.S. Department of Labor, Bureau of Labor Statistics.
20.	1967 Average Hourly Earnings (Same as 20).
Note:	(19) and (20) refer to non-agricultural employment

DESCRIPTION OF THE STRATA

Note on the Stratification

4.10 The combination of areas by strata has a direct bearing on the expected number of areas recommended by DOL program staff that will appear in the sample. To the extent to which recommended areas appear in the same stratum, the number of them expected in the sample goes down. Strata were not built around the recommended areas—partly because there was a total of 26 areas recommended so that with only 10 strata some combination was inevitable, and partly because to attempt to do so would have posed other difficulties in meeting OEO desires for a broad geographic spread of the sample areas. The extent to which the area selection probabilities tend to match program recommendations may be observed from the listing of areas by stratum in Table 14. The average probability for the 26 recommended areas is .032, and for the 16 areas not included in any recommendation is .013.

Characteristics of the Strata

4.11 The characteristics of the 10 strata are described in the following summary:

Stratum	Description
1	Consists of New York City LMA. This is large enough to be a certainty area.
2	Consists of Los Angeles-Long Beach LMA. This combination is about large enough to be a certainty area, and is treated as one.
3-5	Each of these strata is built around an area with a fairly high probability of selection, and represents a special combination.
6-7	These two strata consist of areas in which manufacturing is the largest industry group with regard to employment. The areas all have relatively high average hourly wages in manufacturing, and high median income. They also tend to have relatively low unemployment rates. Stratum 6 consists of areas with more than 5 percent increase in population between 1960 and 1965. Stratum 7 consists of areas with less than 5 percent increase, except for Minneapolis—St. Paul and Seattle. These two were placed in Stratum 7 on the basis that they appeared to match this stratum better than any of the other 9 strata.

Stratum	Description
8-10	These three strata consists of areas in which trade or government is the largest industry group. They are a group of southern and western areas which all have had relatively large percentage increases in population between 1960 and 1965. Strata 8 and 9 represent areas with relatively high average hourly rates in manufacturing, and high median income. Stratum 8 consists of those areas in this category which also have relatively high percentages on non-white population. Stratum 9 represents the balance of the category. Stratum 10 consists of areas with relatively low median income and/or low average hourly rates in manufacturing. Birmingham appears in this stratum as a compromise, even though manufacturing is its largest employer.

The composition of the 10 strata is shown in Table 14.

AREAS AND THEIR SELECTION

4.12 The areas included in the strata are the study areas of the universe recommended in Section II. Geographically contiguous areas which can be administered as a sample unit in the field survey at no greater cost than that for the individual area separately are treated as a sampling unit. These combinations are (a) Dallas-Fort Worth, (b) Los Angeles-Long Beach, and (c) Minneapolis-St. Paul. In addition, if they are selected, Kansas City, Mo. will cover projects in Kansas City, Kans.; St. Louis will cover E.St. Louis; and San Francisco will cover Richmond.

4.13 A sample of 10 areas was selected, one per stratum as follows:

Stratum	Sample Area	Program Recommendations
1	New York City	MDTA, JOBS
2	Los Angeles—Long Beach	MDTA, NC, P
3	Chicago	JOBS, NC, NYC
4	Baltimore	MDTA, NC
5	Detroit	MDTA, P
6	Cincinnati	NYC
7	Pittsburgh	MDTA, NC, P
8	Atlanta	JOBS
9	Dallas—Ft. Worth	NC
10	Birmingham	MDTA, NYC, P

TABLE 14

STRATA FOR A 10-AREA SAMPLE

Stratum	LMA	Measure of Size	DOL Manpower Program Recommendations 1/	Program Organization 2/	
				CEP	S
1	New York City	.100	JOBS, MDTA	I	S
2	Los Angeles—Long Beach	.100	MDTA, NC, P	I	S
3	Chicago	.069	JOBS, NYC, NC	I	X
	St. Louis	.029	JOBS, NC	I	X
		.098			
4	Baltimore	.065	MDTA, NC	I	S
	Newark	.021	MDTA, NYC 3/	I	S
	Philadelphia	.030	MYC	I	S
		.116			
5	Cleveland	.028	P	I	S
	Detroit	.085	MDTA, P	I	S
		.113			
6	Akron	.009		X	S
	Cincinnati	.024	NYC	III	S
	Columbus	.011		II	S
	Dayton	.009		II	X
	Indianapolis	.011		X	S
	Kansas City	.014		II	S
	Louisville	.016		X	X
	Rochester	.005	JOBS, NC, P	II	S
		.099			
7	Boston	.027	MDTA, P 3/	I	S
	Buffalo	.009		II	X
	Jersey City	.022		X	S
	Milwaukee	.010		X	S
	Minneapolis—St. Paul	.017		III	X
	Pittsburgh	.017	MDTA, NC, P	I	X
	Seattle	.006	MDTA, P	II	X
	Toledo	.008		III	X
		.116			

TABLE 14 (Cont)

Stratum	LMA	Measure of Size	DOL Manpower Program Recommendations 1/	Program Organization 2/	
				CEP	S
8	Atlanta	.012	JOBS	I	X
	Houston	.015	MDTA, NYC	I	S
	San Francisco—Oakland	.035	MDTA, P 3/	I	S
	Washington, D.C.	.041		I	X
		.103			
9	Dallas—Ft. Worth	.031	NC	X	S
	Denver	.011	NYC, P 3/	X	X
	Oklahoma City	.004		X	X
	Phoenix	.008		I	S
	Portland	.009		II	X
	San Diego	.011	MDTA, NYC	X	X
		.074			
10	Birmingham	.012	MDTA, NYC, P	I	S
	Memphis	.008		X	S
	Miami	.019	P	II	S
	New Orleans	.008	NYC	I	X
	Norfolk	.014	NYC	II	S
	San Antonio	.026	NYC, P	I	X
	Tampa	.005	NYC	I	X

1/ P denotes an OES recommendation.

2/ S denotes Skill Center. An "X" in either column indicates that there is no CEP or Skill Center, respectively.

3/ One of 5 DOL intensive evaluation cities. In accordance with DOL request, these areas are, in fact, given no chance of selection for the study.

The areas were selected with probability proportional to their measure of size. In this selection, the condition was imposed that not more than one area be selected from a given state. The objective in this was both to avoid clustering of the sample within states and to achieve good geographic spread of the sample. The reason for avoiding clustering of the areas within the state is as follows. The MDTA program is administered through the State Departments of Education, and all of the programs are dependent on the State-run Employment Services. It was thought that these common elements would lead to correlation in the results observed for areas in the same state. If so, this would make the clustering of areas within state inefficient.

4.14 A summary of the program status of the 10 sample areas is given in Table 15.

10-AREA SAMPLE PR

Stratum	LMA	JOBS	Authorized Positions MDTA (Inst)	NYC (O/S)
1	New York City	8,841	14,871	2,940
2	Los Angeles—Long Beach	9,381	8,222	1,329
3	Chicago	6,870	4,808	1,439
4	Baltimore	1,034	1,239	500
5	Detroit	8,502	3,231	753
6	Cincinnati	493	325	570
7	Pittsburgh	1,146	1,696	412
8	Atlanta	883	200	280
9	Dallas—Fort Worth	3,101	1,513	275
10	Birmingham	743	425	300

TABLE 15

PROGRAM STATUS, AS OF FEBRUARY 1969

	Current Non CEP Enrollment				Current CEP Enrollment		
NC	JOBS	MDTA (Inst)	NYC (O/S)	NC	MDTA (Inst)	NYC (O/S)	NC
1,272	2,879	2,376	2,345	1,150	188	---	18
264	3,084	1,618	1,329	225	---	168	342
---	1,609	1,707	1,405	---	1	319	92
---	420	493	633	---	84	115	6
---	1,248	1,409	707	---	211	110	---
---	190	98	560	---	---	---	---
---	739	310	416	---	---	1	74
---	148	302	249	---	144	---	81
120	505	291	260	105		(Non CEP)	
---	194	523	425	---	151	97	94

V. SAMPLING PROGRAM ENROLLEES AND CONTROL CASES

INTRODUCTION

5.1 In this section, some preliminary comments are made on the sampling of program enrollees and control cases. Information for establishing specific sampling plans for the individual programs will be obtained in the course of visits to be made by study teams to each area in advance of the start of the survey, as well as from the pretest for the study to be conducted in New York City. This pretest will also provide some experience for designing the sample of control cases. The discussion in this section deals mainly with the principles to be followed, but also identifies some policy issues concerning which decisions are needed for the final design.

5.2 The study, being based on a sample of enrollees entering the programs during a particular period of time (and a matching sample of control cases), is in fact a study of a particular cohort. The ability to obtain repeated observations on that cohort over time, and to measure the phenomena of change, is a unique contribution of the longitudinal aspect of the study. 1/ However, the target populations, the program designs and operations, and the area-environments may all change over time. Generalization to other populations and different program designs depends upon subject matter expertise to the extent to which the processes studied are not stationary or stable. Such generalization is aided by careful specification and description of the study cohort itself, so that differences between the cohort and other populations of interest can be adequately understood. If the programs can be usefully modeled mathematically, generalization can also be aided by knowledge of changes in the program models. These questions are not considered here.

1/ A useful discussion of the longitudinal approach is given in Dankward Kodlin and Donovan J. Thompson, "An Appraisal of the Longitudinal Approach to Studies of Growth and Development," Monographs of the Society of Research in Child Development, Inc., Vol XXIII, Serial No. 67, No. 1, Child Development Publications, Purdue University, Lafayette, Indiana, 1958.

SAMPLING OF PROGRAM ENROLLEES

5.3 The sampling of program enrollees has two aspects: the selection of projects within the programs, and the sampling of enrollees within the selected projects. By program projects is meant program organizational or operating entities, e.g., contracts in the JOBS program or classes in the MDTA program.

Sampling of Projects

5.4 Generally speaking, the selection of projects within programs will be carried out as a stratified sample appropriate to the program activities in each area. Stratification variables will reflect project size, characteristics of the training occupations, nature of the services provided and/or methods used (including organization and administration) as feasible, and characteristics of enrollees, if these are specified. In order to obtain the desired information, the equivalent of double sampling may be necessary when the number of projects in an area is large. Some special points for the individual programs are noted below.

5.5 Job Corps. It is planned to select the sample of Job Corps enrollees from among those persons recruited from within the 10 study areas, rather than by selecting a sample of Job Corps installations. It is anticipated that any one area will be served by only a limited number of installations, and that the sample of enrollees will be selected from within each one of them. For analyses on an installation basis, it should be noted that with this procedure the probability of selection of a Job Corps installation depends upon the area distribution of its enrollees.

5.6 JOBS. Coverage of enrollees, whether in national or local contracts selected for the study, will be limited to enrollees recruited from within the 10 study areas.

5.7 MDTA. Coverage of enrollees will be limited to members of the "disadvantaged" population. Enrollees for whom training is carried out through individual assignment to nonprogram classes will be sampled separately. Enrollees who are trained in program classes will be sampled through the selection of classes for inclusion in the study.

Sampling of Enrollees Within Projects

5.8 The sampling of enrollees within selected projects will be randomized with controls on characteristics of the enrollees. A definition of "enrollee" for purposes of the study has been proposed to OEO and DOL.

5.9 Controls on the characteristics of the enrollees will be achieved by varying the sampling rates of enrollees by age-sex-race group so as to obtain approximately equal numbers of enrollees in each group to the extent feasible. No controls on other enrollee characteristics will be imposed in the sampling, except for screening with regard to the study definition of the enrollee population to be sampled. In particular, no controls will be imposed to achieve a given distribution of enrollees by referral source.

5.10 No controls will be imposed on the number of enrollees selected in·a particular age-sex-race group from within a given project. A rough control on the total size of the sample from any given project will be achieved by stratifying projects by size and using varying sampling and subsampling rates. The rates will be set on the basis of expected flow of enrollees during the study sampling period so as to spread the sample over that period. It is desirable to extend the study sampling over as long a period of time as possible. This enlarges the size of the cohort, and helps avoid the possibility of being forced to accept an undersirable distribution of enrollees by characteristics. It may also help to avoid biases in the referral of enrollees to the study. The desired average sample per project will be approximately eight enrollees.

SAMPLING OF CONTROLS

Approach

5.11 The control cases will be a sample of the program target populations. A single master control sample will be established covering all the programs, since a given individual can potentially serve as a control for all programs for which he is eligible.

5.12 The sample of control cases will be based on a sample of blocks in designated areas within the sample SMSAs. As a minimum, the designated areas will be defined by updating the poverty areas (major concentrations of poverty) established for OEO by the Bureau of the Census. Within selected blocks, a sample of addresses will be selected and screening interviews conducted to identify individuals who would be eligible for the five manpower training programs and, hence, are potential controls. If it is found to be feasible within the study resources, it would be desirable to avoid this restriction and to extend the sampling over the study area. This question will be explored further with the study contractor. Objective definitions of program eligibility which LS&R proposes to use for this purpose have been approved by OEO and DOL program staffs, and are the basis for design of the screening questionnaire.

Scheduling of Screening

5.13 If the screening interviews follow the sampling of program enrollees, the total amount of such interviewing can be minimized, since the characteristics of the desired control sample is known. The New York City pretest will provide information as to the feasibility of such an arrangement of the field operations. If the screening precedes the sampling of program enrollees, the lag between the two should be minimized. This is so because the longer the lag, the greater the opportunity for screenees designated for the sample to move before they are selected. In such instances it is necessary to follow-up the selected control case to his new address, rather than substituting another control for him, to avoid bias in the control sample. (Without such follow-up the more mobile control cases would tend to be under-represented in the sample.) Similarly to avoid bias, the eligibility of screenees for the study should be based on their characteristics at the time of screening and not at the time of follow-up for the Wave I interview (these latter are known only for the screenees selected as control cases).

Matching in Selection of Controls

5.14 The actual selection of controls will be carried out as an office operation to avoid the bias thought to be inherent in interviewer selection. Only one control case will be selected from any given household. Rough control on the total size of the control sample selected from any one block will be achieved by the sampling technique.

5.15 Control cases will be sampled with stratified matching of program enrollees. The question of variables to be used for matching beyond age-sex-race is open. One additional variable for matching being considered is geographic area of residence. However, it is not planned to use more than a few variables for matching. It is expected to achieve most of the statistical efficiency of more matching variables by suitable techniques in the analysis of the data. This is the more desirable approach when possible since the multiplication of matching variables would drive up the amount of screening to obtain the control cases and, hence, screening costs. Moreover, it avoids the possible risk of obtaining an unusual sample of the target population.

5.16 Because there will be more program enrollees than control cases in the study sample, a program "running mate" for each control case will be designated. The successive study interviews with each control case will then be scheduled to take place at approximately the same points in time as those for his running mate. The program running mates for the controls will be selected as a probability sample.

SOME POLICY ISSUES

Enrollee Population to Be Sampled

5.17 Application of Program Eligibility Criteria. There is evidence to suggest that programs enroll some individuals who would be ineligible under a strict application of program eligibility criteria. It is tentatively proposed to select the enrollee sample only from among those enrollees who meet program eligibility criteria. This decision is subject to review in the light of further evidence, including the New York City pretest. In any event, the final criteria for inclusion in the study population must be applied uniformly in both the sampling of enrollees and the sampling of control cases.

5.18 Restrictions on Enrollee Population Sampled. Questions have been raised as to whether the enrollee population to be sampled for the study should be restricted, for example by excluding older or disabled persons. The resolution of this problem should be based on consideration of the benefits for the study to be gained by the restrictions.

5.19 Central City-Balance of SMSA Distribution. The universe of study areas has been defined in terms of SMSAs (LMAs). It appears that program activities are heavily concentrated within the central cities of SMSAs. The study has the option to sample program projects over the entire SMSA (LMA) or only within the central city. The balance of the SMSAs outside of the central cities accounts for a substantial proportion of the poverty population from which the program target populations would be drawn, according to OEO-DOL estimates. Further, these estimates suggest that the racial composition of the target population may be very different (more heavily nonwhite) in the central city than in the balance of the SMSA.

5.20 It is tentatively proposed to sample program activities over the entire SMSA. Correspondingly the screening sample for control cases would also be selected over the entire SMSA. This decision should be reviewed on the basis of its expected impact on the study data collection costs.

Target Population Sampling

5.21 The sample of the target population to serve as program controls will be selected from eligible individuals identified by screening interviews of a sample of households (addressees). Experience shows that this type of screening tends to undercover young males, particular nonwhites. The impact on the study would be that the control sample for this population group would be biased due to under-representation of the types of individuals subject to undercoverage. (The impact on study costs would be that more screening would be required to provide a given number of control cases.) In some studies, an effort has been made to obtain representation of persons likely to be missed in the screening interview by selecting individuals found on the street or in places such as bars, pool halls and barber shops. Such "uncoventional" techniques have the problem that the composition of the sample compared with the population for the different types of individuals in unknown, and may itself be biased.

5.22 It is speculated that the undercoverage of young males in the screening interviews might be on the order of 10 percent; and that the labor force status of those missed would not be far different from that of those covered in the screening interviews. On this basis, it is tentatively proposed to accept the bias of the screening interviews in representing the target population rather than the uncertainties of the "uncoventional" techniques. It is also proposed to monitor the results of the screening on a current basis. If the assumptions of the study design appear to have been too optimistic, corrective steps will be instituted.

VI. SOME OPPORTUNITIES IN THE RESEARCH DESIGN

INTRODUCTION

6.1 There are a number of opportunities in the research design for strengthening the study data and the span of inference from the study. Some of these are briefly reviewed in this section for future reference.

PROGRAM COVERAGE

6.2 If the New Careers program is not included in the study, and the funds are available within the study budget, it would be of interest to substitute another program for it. Two programs which might be considered are:

- MDTA, On-the-Job Training component
- NYC, In-School component .

The size and area distribution of each of these is such that they would cost less to cover than the New Careers program with the sampling error specifications achieved for estimates for the other study programs. MDTA (OJT) would be a natural comparison for JOBS (Contract) and MDTA (Inst.), and would be more feasible than Vocational Education within the study. The NYC (I/S) would add an additional program for the under 22 age group.

PROGRAM ORGANIZATION

6.3 In the study special aspects of program organization such as CEP's or skill centers are expected to be covered only in proportion to their role in the training programs. For example, if 10 percent of the enrollment in a program were funded through CEP's in the study areas then it would be expected that 10 percent of the study sample would be CEP referrals. It would be possible to sample enrollees at differential rates to strengthen the study estimates for evaluation of special aspects of program organization or to supplement the planned study sample for this purpose. The first approach would have the effect of increasing the sampling errors of the overall study estimates; the second would increase the study costs.

TARGET POPULATION

6.4 If funds are available, it would be exceedingly valuable to select subsamples of control cases during the course of the follow-up period and to assign them to the outreach and recruitment components of the programs, continuing to follow them whether or not they actually enroll. The resulting data would, first of all, provide measures comparable to an experimental design rather than an observational study, as is now provided by the study design. (In an experimental design both the program and control cases should represent random samples of the same population.) Second, the data would provide additional insight as to the nature of the target population, and the extent to which the program target populations are, in fact, accessible to current outreach and recruitment techniques. If this were to be done, the initial control sample would need to be enlarged.

6.5 From the point of view of strengthening the study data for future program planning, it would be of interest to follow samples of the target populations of the programs which are not represented in current program operations. These data would be useful in supporting analyses of the potential impacts of changes in the de facto target populations of the programs. Some background information for such analyses will, of course, be available in any event by including in the study analyses data from those units in which screening identified program eligibles but from which no control case was selected.

VII. CONCLUSIONS AND RECOMMENDATIONS

7.1 In this report, the analyses carried out for the design of a sample for the longitudinal evaluation study of five major U.S. Government manpower training programs have been described. A major focus of these analyses was the OEO-DOL request for recommendations as to the adequacy of their preliminary specifications for a study sample of 10 areas and 10,000 persons. The analyses were carried out with a view to programmatic uses of the data to come from the study, rather than just a "benefit-cost" analysis for the particular enrollees in the programs at the time of the study. Programmatic uses of the data are considered to require primarily estimates of program impacts, and of the sampling errors of such estimates, rather than tests of significance. Conclusions and recommendations from the sample design analyses are summarized in this section of the report.

PRELIMINARY OEO-DOL SAMPLE SPECIFICATIONS

7.2 The preliminary specifications for the study sample established by OEO and DOL can be expected to provide estimates and analyses within programmatically useful limits of sampling error for four programs:

Job Corps
JOBS (Contract)
MDTA (Inst.)
NYC (O/S).

7.3 It is recommended that the New Careers program be dropped from the study or be budgeted separately. The area distribution of New Careers is markedly different than that of the other programs. To include it without separate budgeting, and correspondingly reduce the sample size for the other programs, would increase the sampling errors of estimates and analyses for the other programs without providing estimates for New Careers on a basis comparable to those for the other programs. If the New Careers program is budgeted separately, some additional areas should be selected for the study to strengthen the estimates and analyses for New Careers. Also, a longer initial period for accumulating the desired sample of enrollees should be planned for than in the case of the other programs.

UNIVERSE OF STUDY AREAS

7.4 It is recommended that the universe of areas from which the study areas will be selected be taken as the Labor Market Areas corresponding to SMSAs of 500,000 or more population in 1960 with central city of 250,000 or more. This universe includes 43 of the 46 JOBS cities in conterminous U.S. The JOBS cities not included are El Paso, Omaha, and Tulsa. Excluded outside conterminous U.S. are Honolulu and 3 JOBS cities in Puerto Rico. The establishing of a universe of study areas for sampling helps clarify the universe for which statistically-based inferences from the study will be possible; and the universe for which, since it was not sampled, inference will depend on subject-matter expertise. From the definition of the universe of study areas, it is clear that the study will be an urban, not a rural, one. Such rural places as may be represented in the study sample will be of the type found in the SMSAs of large cities.

STUDY SAMPLE DESIGN

Sample of Areas

7.5 It is recommended that the 10 areas in which the study will be conducted be selected as a probability sample. Since individuals will be designated for the study on enrollment, as they are referred by the programs, the fact that they are the subjects of evaluation will be known to program staffs. This fact could be reflected in special selection and/or treatment of enrollees. If favorable outcomes for enrollees are observed in a probability sample of areas, the inference is that what was accomplished in the study areas can be accomplished elsewhere. If the study areas were chosen subjectively it would be difficult to disprove the argument that this inference should not be drawn.

7.6 A recommended design for the sampling of areas is described. The design is intended to be reasonably efficient for the various study objectives and programs.

Allocation of Sample of Individuals

7.7 It is recommended that the initial sample of individuals to be included in the study consist of 7,500 program enrollees and 3,500 control cases, to be allocated equally by program group and by area to the extent feasible. The cost of this sample of 11,000 individuals is believed to be equivalent to that intended under the preliminary OEO-DOL specification for 10,000 individuals to be included in the study. It would provide a total initial sample of 1,250 enrollees in each of the Jobs Corps and NYC (O/S) programs, and in each of the two age group's (under 22, and 22 and over) of the JOBS and MDTA (Inst.) programs, and

1,750 control cases for each of the two age groups. If the sample attrition is held to 20 percent over the life of the study this would provide a final sample of 1,000 individuals per group for the analysis. This sample allocation can be expected to permit estimates of post-program changes in average annual earnings per enrollee and benefit-cost ratios, within programmatically useful limits of sampling error.

7.8 It is recommended that the sample of enrollees for each program be allocated equally among the four race-sex groups to the extent feasible. This allocation is intended to provide estimates of post-program changes in average annual earnings per enrollee by race-sex group within useful limits of sampling error for each, as well as to provide a basis for further analyses of the impacts of program components and the factors in success or failure for each of the age-sex groups. From the data available for analysis on this point, it appears that an allocation of the sample by race and sex proportional to the mix of program enrollees that might be encountered during the study sampling period would be likely to provide poor estimates for whites and, to a lesser extent, for females.

SPECULATED SAMPLING ERRORS

Post-Program Changes in Income

7.9 Estimates of sampling errors to be expected in estimating post-program changes in average annual income per program enrollee, compared with their controls, were constructed using data from earlier evaluation studies made available by OEO and DOL. Because of the limitations in the available data, these estimates of expected sampling errors are referred to in the report as "speculated sampling errors." Despite the qualifications and uncertainties that are attached to the speculated sampling errors, it is felt that they provide a reasonable indication of the level of sampling errors to be expected in independent estimates from the study. An illustrative summary of the results of the sampling error analysis is given in the following tabulation for estimates of post-program change in average annual earnings for all enrollees in a given program compared with their controls.

If the average annual earnings per control case are	The chances are about 2 out of 3 that the difference between the estimated change and the change in annual earnings derived from a study based on all enrollees in the program would be less than	
	Estimated change for study areas	National projection of change to all areas
$ 500	$ 40-$ 50	$ 60-$ 75
$1,000	$ 65-$ 75	$100-$110
$2,000	$110-$115	$150-$160

The chances are about 19 in 20 that the difference between the change estimated from the study sample and that which would be found from a study of all enrollees would be less than twice the limits given in this tabulation. In Section III of this report a more detailed analysis of the speculated sampling errors for estimates of post-program changes in earnings by age-sex-race group, and for estimated benefit-cost ratios, is presented. Also presented are estimates of true changes in post-program earnings for which the odds (or probability) that the change would be detected statistically in the study are suitably high.

7.10 The criterion of programmatically useful sampling errors adopted in this report is that the uncertainty in estimated post-program changes in annual earnings will be within \$300-\$400. Put another way, there should be high assurance that if changes of this order of magnitude exist they will be detected statistically in the study. It is considered that changes on the order of \$50-\$100 a year are of questionable significance either for program planning or for potential program enrollees. Moreover, there is not sufficient accuracy in technique for measuring total annual income to afford much confidence in differences on the order of \$50-\$100 a year.

Benefit-Cost Ratios

7.11 It is speculated that the sampling errors of benefit-cost ratios estimated from the study might be as high as 10 to 20 percent for ratios estimated for the study areas, and 15 to 25 percent for national projections of ratios to all areas, under some combination of benefits and costs.

7.12 Although the speculated sampling errors appear to be high compared to the specifications for precision ordinarily met in survey studies, it is suggested that they are useful for purposes of the study analysis. For example, a 20 percent sampling error in observed benefit-cost ratios would have the implications summarized in the following tabulation.

If the observed benefit-cost ratio is	The chances are about 2 out of 3 that the difference between the observed ratio and the benefit-cost ratio from a study of all program enrollees would be less than
1	0.2
2	0.4
5	1.0
10	2.0

The uncertainty due to sampling which is illustrated by this tabulation is relatively small compared to that arising from other sources of uncertainty which affect the estimated benefit-cost ratios. Among such factors which affect the level of an estimated benefit-cost ratio are the assumptions as to the patterns of benefits to be projected for time periods not directly observed, the length of the time horizon over which benefits are projected, and the choice of an appropriate rate for discounting future benefits.

SAMPLING OF PROGRAM ENROLLEES AND CONTROL CASES

Comments on the Sampling Approach

7.13 The study, being based on a sample of enrollees entering the programs during a particular period of time (and a matching sample of control cases) is, in fact, a study of a particular cohort. Generalization to other populations and different program designs depends upon subject matter expertise to the extent to which the processes studied are not stationary or stable. Such generalization is aided by careful specification and description of the study cohort itself, so that differences between the cohort and other populations of interest can be adequately understood.

7.14 It is recommended that the program enrollees and control cases for the study be selected as probability samples. In the report some preliminary comments are made on the principles to be followed in sampling of program enrollees and control cases in the study. It is recommended that the study sampling be extended over as long a time period as feasible. This enlarges the size of the cohort, and helps avoid the possibility of being forced to accept an undesirable distribution of enrollees by characteristics.

Some Issues for Decision

7.15 Several issues for decision are identified in the report:

a. Programs enroll some individuals who would be ineligible under a strict application of program criteria. It is tentatively proposed to select the enrollee sample only from those individuals who meet program eligibility criteria. This is subject to review after the New York City pretest.

b. Questions have been raised as to whether the enrollee population to be sampled for the study should be restricted, for example, by excluding older or disabled persons. These are to be resolved by consideration of the benefits for the study to be gained by the restrictions.

c. It is tentatively proposed in each study area, to sample program activities over the entire SMSA. This is subject to review on the basis of the expected impact on program costs.

SOME OPPORTUNITIES IN THE RESEARCH DESIGN

7.16 Some opportunities in the research design for strengthening the study data and the span of inference from the study if funds are available are identified in the report.

APPENDIX A

TABLES FOR SECTION II

TABLE A.1

50 JOBS CITIES

State	City
Alabama	Birmingham
Arizona	Phoenix
California	Long Beach Los Angeles Oakland San Diego San Francisco
Colorado	Denver
District of Columbia	Washington
Florida	Miami Tampa
Georgia	Atlanta
Hawaii	Honolulu
Illinois	Chicago
Indiana	Indianapolis
Kentucky	Louisville
Louisiana	New Orleans
Maryland	Baltimore
Massachusetts	Boston
Michigan	Detroit
Minnesota	Minneapolis St. Paul
Missouri	Kansas City St. Louis
Nebraska	Omaha
New Jersey	Jersey City Newark
New York	Buffalo New York Rochester

TABLE A.1(Cont)

State	City	State	City
Ohio	Akron Cincinnati Cleveland Columbus Dayton Toledo	Texas	Dallas El Paso Fort Worth Houston San Antonio
		Virginia	Norfolk
Oklahoma	Oklahoma City Tulsa	Washington	Seattle
Oregon	Portland	Wisconsin	Milwaukee
Pennsylvania	Philadelphia Pittsburgh		
Tennessee	Memphis		

TABLE A.2
REGIONAL OFFICES FOR HEW, LABOR, HUD, OEO, SBA

Region	Regional Area	Location Regional Office	Jurisdiction		
I	New England	Boston	Conn.	R.I.	
			Maine	Mass.	
			N.H.	Vt.	
II	Northeast	New York	N.Y.	Pr.	
			N.J.	Vt.	
III	Mid Atlantic	Philadelphia	Del.	Md.	Va.
			D.C.	N.C.	W.Va.
			Ky.	Pa.	
IV	Southeast	Atlanta	Ala.	Miss.	
			Fla.	S.C.	
			Ga.	Tenn.	
V	Midwest	Chicago	Ill.	Minn.	
			Ind.	Ohio	
			Mich.	Wisc.	
VI	Southwest	Dallas-Fort Worth	Ark.	Okla.	
			La.	Tex.	
			N.Mex.		
VII	Mountain Plains	Denver	Colo.	Mo.	So.Dak.
			Idaho	Mont.	Utah
			Iowa	Neb.	Wyo.
			Kan.	N.Dak.	
VIII	Far West	San Francisco	Alaska	Hawaii	Wash.
			Ariz.	Nev.	Guam
			Cal.	Ore.	

TABLE A.3

DOL MANPOWER PROGRAMS REPORTED IN STATUS REPORT OF PROJECTS ACTIVE AS OF FEBRUARY 1969

Training Programs:
- MDTA Institutional
- MDTA On-the-Job
- MDTA Apprenticeship
  - Outreach
- JOBS
- New Careers

Work-Experience Programs

Neighborhood Youth Corps
- In-School
- Out-of-School
- Summer

Operation Mainstream

Special Programs:
- Experimental and Demonstration
- Special Impact

Concentrated Employment Program

Work Incentive Program

APPENDIX B

BIBLIOGRAPHIC NOTE FOR SECTION III

B.1 The literature on the design and use of sample survey data for "analytical" purposes is not extensive; analysis of variance models to facilitate the design and analysis of such surveys have been employed by J. Sedransk in a series of papers, although he has not specifically considered the question of programmatic applications raised here. See, for example, "Analytical Surveys with Cluster Sampling," J. Roy Statist. Soc., Series B, 27 (1965), 264-278, in which a mixed model is used; and "Designing Some Multi-Factor Analytical Studies," J. Amer. Statist. Assoc., 62 (1967), 1121-1139, in which a fixed effects model is used. These papers assume that the goal of an analytical survey is to compare the means of different domains of study (sub-populations). In "Planning Some Two-Factor Comparative Surveys," J. Amer. Statist. Assoc., 64 (1969), 560-573, Sedransk and Booth consider the design of a survey to make overall comparisons between classes for one classification of a two (or more)- way classification. The problem considered is related to that of this study. Frank Yates has a discussion of methods for constructing estimates of the means of classes for one classification or a two-way classification and of the limitations of sample survey data for analytical purposes, because they are observational rather than experimental data, in his "Sampling Methods for Censuses and Surveys," 3rd Ed., Sec. 5.23-5.25 and 9.6-9.10, Hafner Publishing Company, New York, 1960. Some useful comments on problems of design and interpretation are given in N. Keyfitz, "A Factorial Arrangement of Comparisons of Family Size," J. Amer. Sociol., 58 (1953), 470-480; and W.G. Cochran, Matching in Analytical Studies," Amer. J. Pub. Health, 43 (1953), 684-691. Models for the analysis of variance with infinite and finite population sampling, and questions of inference from a sampled universe to other universes, are discussed in J. Cornfield and J.W. Tukey, "Average Values of Mean Squares in Factorials," Ann. Math. Statist., 27(1956), 907-949; M.B. Wilk and O. Kempthorn. "Some Aspects of the Analysis of Factorial Experiments in a Completely

Randomized Design," Ann. Math. Statist. 27 (1956), 950-985; and P.J. McCarthy (op. cit). Statistical methodology appropriate for the analysis of data from complex surveys, in the sense of being "exact," is reviewed and developed by P.J. McCarthy in two publications: "Replication, An Approach to the Analysis of Data from Complex Surveys," National Center for Health Statistics, Series 2, No. 14 (April 1966); and "Pseudoreplication, Further Evaluation and Application of the Balanced Half-Sample Technique," National Center for Health Statistics, Series 2, No. 31 (January 1969); U.S. Government Printing Office, Washington, D.C.

APPENDIX C

REGRESSION ANALYSIS FOR ESTIMATES BY ANALYSIS GROUPS

C.1 In the study, estimates will be prepared for a variety of analysis groups to provide measures of the effect on program impact of a number of variables of interest; for example, race and sex of the trainee. One alternative to separate estimates for each such analysis group, based on the samples from the groups in the study, is to derive such estimates from a regression equation based on the entire sample with dummy variates representing the groups. This approach is one typically used in econometric studies. [1] When there is a large number of factors of interest, methods such as this are of great potential value in permitting detailed analysis without the necessity for corresponding proportionate increases in the sample size required.

C.2 The objective of this Appendix is to briefly explore the potential contribution of this approach to the study analysis. The analysis compares the sampling errors of independent estimates with those of regression estimates. A simplified statistical model is used in order to exhibit the structure of the results.

---

[1] For a general discussion of this approach in the context of regression analysis, see, for example, Arthur S. Goldberger, Econometric Theory, John Wiley & Sons, Inc., New York, 1964; E. Malinvaud, Statistical Methods of Econometrics, North-Holland Publishing Company, Amsterdam, 1966. For a discussion in the context of the analysis of variance of multiple classifications, see, for example, Oscar Kempthorne, The Design and Analysis of Experiments, John Wiley & Sons, Inc., New York, 1952; H. Scheffe, The Analysis of Variance, John Wiley & Sons, Inc., New York, 1959.

The Model

C.3 The model assumes that separate regressions are computed for each study area. The regression analysis is to provide estimates of the expected value of some variable of interest, say y, for groups defined by combinations of the analysis factors. For example, y might be annual post-program earnings. To be specific, assume that estimates are to be made for each of two race groups by sex, and that within each of the four race-sex groups a simple random sample of n individuals is selected [2] for each of which an observation on the variable of interest is obtained.

C.4 Let

y_{ijk} = observed value of y for the k-th sample individual in race group i and sex j; $i = 1,2$; $j = 1,2$; $k = 1,2, \ldots, n$.

The y_{ijk} are assumed to be normally and independently distributed with means μ_{ij} and common variance σ^2.

Independent Estimates by Race-Sex Group

C.5 If independent estimates are made by race-sex group, the minimum variance unbiased estimator of μ_{ij} is $\bar{y}_{ij}$, the mean of the n sample observations for individuals in race group i and sex j. The variance of each of the estimates $\bar{y}_{ij}$ is

$$\sigma^2_{\bar{y}_{ij}} = \frac{\sigma^2}{n}. \qquad (C.1)$$

Regression Estimates by Race-Sex Group

C.6 The regression estimates will differ according to whether or not there is interaction between the effects of the two variables, race and sex.

C.7 Interaction Assumed to Be Absent. If there is no interaction, it is assumed that

$$\mu_{ij} = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} \qquad (C.2)$$

where X_{1ij} and X_{2ij} are dummy variates for race and sex, respectively, with

$$X_{1ij} = \begin{cases} +1 & \text{for an individual in race group 1} \\ -1 & \text{for an individual in race group 2} \end{cases}$$

$$X_{2ij} = \begin{cases} +1 & \text{for an individual of sex 1} \\ -1 & \text{for an individual of sex 2} \end{cases}$$

---

[2] From an infinite population of such individuals.

The minimum variance unbiased estimator of μ_{ij} is

$$\bar{y}_{ij\,(\text{reg.})} = b_0 + b_1 X_{1ij} + b_2 X_{2ij} \tag{C.3}$$

where the b's are the estimates of the corresponding β's defined by the conventional least squares estimators. Since the number of observations in each of the four race-sex groups is the same, the dummy variates are orthogonal and

$$b_0 = \bar{y}_{..} \tag{C.4}$$

$$b_1 = \tfrac{1}{2}\,(\bar{y}_{1.} - \bar{y}_{2.}) \tag{C.5}$$

$$b_2 = \tfrac{1}{2}\,(\bar{y}_{.1}1 - \bar{y}_{.2}) \tag{C.6}$$

where $\bar{y}_{i.}$ is the mean of y for race group i, $\bar{y}_{.j}$ is the mean of y for sex j, and $\bar{y}_{..}$ is the overall mean. Thus,

$$\bar{y}_{ij\,(\text{reg})} = \bar{y}_{..} + \tfrac{1}{2}\,(\bar{y}_{1.} - \bar{y}_{2.})\,X_{1ij} + \tfrac{1}{2}\,(\bar{y}_{.1} - \bar{y}_{.2})\,X_{2ij}\,. \tag{C.7}$$

These estimates may be written explicitly as follows:

$$\left.\begin{aligned}
\bar{y}_{11}\,(\text{reg}) &= \bar{y} + \tfrac{1}{2}\,(\bar{y}_{1.} - \bar{y}_{2.}) + \tfrac{1}{2}\,(\bar{y}_{.1} - \bar{y}_{.2}) \\
&= \tfrac{1}{4}\,(3\,\bar{y}_{11} + \bar{y}_{12} + \bar{y}_{21} - \bar{y}_{22}) \\[1em]
\bar{y}_{12}\,(\text{reg}) &= \bar{y} + \tfrac{1}{2}\,(\bar{y}_{1.} - \bar{y}_{2.}) - \tfrac{1}{2}\,(\bar{y}_{.1} - \bar{y}_{.2}) \\
&= \tfrac{1}{4}\,(\bar{y}_{11} + 3\,\bar{y}_{12} - \bar{y}_{21} + \bar{y}_{22}) \\
\bar{y}_{21}\,(\text{reg}) &= \bar{y} - \tfrac{1}{2}\,(\bar{y}_{1.} - \bar{y}_{2.}) + \tfrac{1}{2}\,(\bar{y}_{.1} - \bar{y}_{.2}) \\
&= \tfrac{1}{4}\,(\bar{y}_{11} - \bar{y}_{12} + 3\,\bar{y}_{21} + \bar{y}_{22}) \\
\bar{y}_{22}(\text{reg}) &= \bar{y} - \tfrac{1}{2}\,(\bar{y}_{1.} - \bar{y}_{2.}) - \tfrac{1}{2}\,(\bar{y}_{.1} - \bar{y}_{.2}) \\
&= \tfrac{1}{4}\,(-\,\bar{y}_{11} + \bar{y}_{12} + \bar{y}_{21} + 3\,\bar{y}_{22})
\end{aligned}\right\} \tag{C.8}$$

The variance of each of the estimates $\bar{y}_{ij\,(reg)}$ is

$$\sigma^2_{\bar{y}_{ij}\,(reg)} = 0.75\,\frac{\sigma^2}{n} = 0.75\,\sigma^2_{\bar{y}_{ij}} \qquad (C.9)$$

and

$$\sigma_{\bar{y}_{ij}\,(reg)} = 0.87\,\sigma_{\bar{y}_{ij}}\,. \qquad (C.10)$$

In the general case of r factors, each having two levels, the variance of $\bar{y}_{ijk\ldots(reg)}$ is

$$\sigma^2_{\bar{y}_{ijk\ldots(reg)}} = \frac{1+r}{2^r}\,\frac{\sigma^2}{n} = \frac{1+r}{2^r}\,\sigma^2_{\bar{y}_{ijk\ldots}} \qquad (C.11)$$

where n is the number of observations per cell. If there are a total of m observations altogether $n = m/2^r$ and

$$\sigma^2_{\bar{y}_{ijk\ldots}} = 2^r\,\frac{\sigma^2}{m} \qquad (C.12)$$

while

$$\sigma^2_{\bar{y}_{ijk\ldots(reg)}} = (1+r)\,\frac{\sigma^2}{m}\,. \qquad (C.13)$$

C.8 Interaction Assumed to Be Present. If there is interaction between the effects of the two variables race and sex, it is assumed that

$$\mu_{ij} = \beta_0 + \beta_1\,x_{1ij} + \beta_2\,x_{2ij} + \beta_3\,x_{3ij} \qquad (C.14)$$

where x_{1ij} and x_{2ij} are defined as above and [3]

$$x_{3ij} = x_{1ij} \cdot x_{2ij}\,. \qquad (C.15)$$

---

[3] See, for example, Einar Hardin and Michael E. Borus, Economic Benefits and Costs of Retraining Courses in Michigan, School of Labor and Industrial Relations, College of Social Science, Michigan State University, June 1967 (Processed).

The minimum variance unbiased estimator of u_{ij} in this case is

$$\bar{y}^*_{ij\,(\text{reg})} = b_0 + b_1 x_{1ij} + b_2 x_{2ij} + b_3 x_{3ij} \qquad (C.16)$$

where b_0, b_1, and b_2 are defined as in Equations (C.4) to (C.6)

and

$$b_3 = \tfrac{1}{4} (\bar{y}_{11} - \bar{y}_{12} - \bar{y}_{21} + \bar{y}_{22}) \,. \qquad (C.17)$$

Thus the coefficient b_3 may be interpreted as representing the difference between the two race groups, of the difference between sex within race, or the difference between the two sex groups of the difference between race within sex.

$$\begin{aligned} \bar{y}^*_{ij\,(\text{reg})} &= \bar{y}_{..} + \tfrac{1}{2} (\bar{y}_{11} - \bar{y}_{2.}) x_{1ij} + \tfrac{1}{2} (\bar{y}_{.1} - \bar{y}_{12}) x_{2ij} \\ &\quad + \tfrac{1}{4} (\bar{y}_{11} - \bar{y}_{12} - \bar{y}_{21} + \bar{y}_{22}) x_{3ij} \,. \end{aligned} \qquad (C.18)$$

If these estimates are computed explicitly, it is found

that

$$\bar{y}^*_{ij\,(\text{reg})} = \bar{y}_{ij} \qquad (C.19)$$

as might have been anticipated, so that there is no increase in efficiency compared with independent estimates by race-sex group.

APPENDIX D

SAMPLE DESIGN ANALYSIS

D.1 This Appendix presents supplementary details of the discussion in Sections III, IV, and V.

OPTIMIZATION ANALYSIS

D.2 The following details supplement paragraphs 3.11-3.20.

Optimum Selection Probabilities and Sampling and Subsampling Rates

D.3 With the sampling model and estimator specified in paragraphs 3.14 - 3.15, the sampling variance of r with the usual approximation for a ratio estimator is found to be [1]

$$\sigma_r^2 = \frac{1}{qy^2} \sum_{h=1}^{Q} \frac{N_h^2 Z_h^2}{P_h} + \frac{1}{qy^2} \sum_{h=1}^{Q} \frac{N_h^2}{P_h} \frac{\sigma_{wh}^2}{n_h} \qquad (D.1)$$

where

$$Z_h = (\bar{X}_h - R\bar{Y}_h)^2$$

$$\bar{X}_h = E(\bar{x}_h/h) \quad \bar{Y}_h = E(\bar{y}_h/h) \quad R = \bar{X}_h/\bar{Y}_h$$

$$\sigma_{wh}^2 = \sigma_{hx}^2 - 2R\sigma_{hxy} + R^2 \sigma_{hy}^2$$

$$\sigma_{hx}^2 = E\left[(x_{hi} - \bar{X}_h)^2 / h\right] \quad \sigma_{hy}^2 = \left[(y_{hi} - \bar{Y}_h)^2 / h\right]$$

---

[1] The methods for deriving these results parallel those presented in M.H. Hansen, W.N. Hurwitz, and W.G. Madow, Sample Survey Methods and Theory, Vol II, Theory, John Wiley & Sons, Inc., New York, 1953, Chapter 9; and W.G. Cochran, Sampling Techniques (Second Edition), John Wiley & Sons, Inc., New York, 1963, Chapter 11.

$$\sigma_{hxy} = E\left[\left[(x_{hi} - \bar{X}_h)(y_{hi} - \bar{Y}_h)/h\right]\right]$$

$$Y = \sum_{h=1}^{Q} Y_h \qquad Y_h = N_h \bar{Y}_h$$

and E(-/h) denotes the conditional expected value of the variate inside the parentheses for h fixed, i.e., for a given area. Since the sample size will vary depending upon the particular areas falling in the sample, the cost used for the optimization is the expected total cost (average over all samples)

$$c_1 q + c_2 q \sum_{h=1}^{Q} P_h n_h \tag{D.2}$$

If the sample is such that the estimates x' and y' are self-weighting

$$\frac{n_h}{N_h} P_h = k \qquad \sum_{h=1}^{Q} P_h = 1$$

and the variance (D.1) may be written

$$\sigma_r^2 = \frac{1}{qY^2} \sum_{h=1}^{Q} \frac{N_h^2 Z_h^2}{P_h} + \frac{N\sigma_w^2}{qkY^2} \tag{D.3}$$

where $\sigma_w^2 = \sum_{h=1}^{Q} N_h \sigma_h^2 / N \qquad N = \sum_{h=1}^{Q} N_h,$

and the cost (D.2)

$$c_1 q + c_2 qkN \tag{D.4}$$

To carry out the optimization set up the Lagrangian ψ:

$$\psi = \sigma_r^2 + \lambda_1 (c_1 q + c_2 qkN - c) + \lambda_2 (\Sigma P_h - 1) \tag{D.5}$$

where c is the fixed total cost. Then the solution of the equations $\partial\psi/\partial q = 0$, $\partial\psi/\partial P_h = 0$, $\partial\psi/\partial k = 0$, $\partial\psi/\partial\lambda_1 = 0$, $\partial\psi/\partial\lambda_2 = 0$ will give the values of P_h, q, and k which minimize the variance of r subject to fixed cost. The result for P_h is:

$$P_h = \frac{\sqrt{N_h^2 Z_h^2}}{\sum_h \sqrt{N_h^2 Z_h^2}} = \frac{\sqrt{N_h^2 \delta_h \sigma^2}}{\sum_h \sqrt{N_h^2 \delta_h \sigma^2}} \qquad \text{(D.6)}$$

where $Z_h^2 = \delta_h \sigma^2$ (D.7)

and $\sigma^2 = \sigma_B^2 + \sigma_w^2$ (D.8)

with $\sigma_B^2 = \frac{1}{Q} \sum_{h=1}^{Q} Z_h^2 .$ (D.9)

Let $\delta = \frac{1}{Q} \sum_h \delta_h$ (D.10)

then, $\sigma_B^2 = \delta \sigma^2 \qquad \sigma_w^2 = (1-\delta)\sigma^2 .$ (D.11)

The result for k is

$$k = \sqrt{\frac{c_1}{c_2}} \frac{\sqrt{1-\delta}}{\sum \sqrt{N_h^2 \delta_h}} \qquad \text{(D.12)}$$

so that the optimum n_h are

$$n_h = \sqrt{\frac{c_1}{c_2}} \sqrt{\frac{1-\delta}{\delta_h}} . \qquad \text{(D.13)}$$

If the δ_h do not vary greatly, so that they may be replaced by their average as an approximation, the optimum values for k and n_h become

$$\left.\begin{aligned} k &\doteq \frac{1}{N} \sqrt{\frac{c_1}{c_2}} \sqrt{\frac{1-\delta}{\delta}} \\ n_h &\doteq \sqrt{\frac{c_1}{c_2}} \sqrt{\frac{1-\delta}{\delta}} \end{aligned}\right\} \qquad \text{(D.14)}$$

(This condition is also that for a self-weighting sample to be optimum.)

D.4 In applying for results of this analysis, the optimum values for the P_h and k are determined. The optimum values for q, the number of areas to select, is then determined by the total fixed cost from the formula

$$q = \frac{c}{c_1 + c_2 kN} \quad . \tag{D.15}$$

The structure of the optimum thus parallels the intuitive approach of determining the sample required per area for the analysis contemplated, with the number of areas to be included in the study then being fixed by the available funds.

Implications of Compromise Selection Probabilities

D.5 The selection probabilities arrived at for the study sample, as described in paragraph 3.19, are not the optimum probabilities that would be indicated by Equation (D.6) for each of the programs considered individually, but a compromise between them. For an arbitrary set of probabilities, the values of k and n_h corresponding to those in (D.14) are

$$\begin{aligned} k &\doteq \sqrt{\frac{c_1}{c_2}}\sqrt{\frac{1-\delta}{\delta}} \; \frac{1}{\sqrt{\sum N_h^2/P_h}} \\ n_h &\doteq \sqrt{\frac{c_1}{c_2}}\sqrt{\frac{1-\delta}{\delta}} \; \frac{N_h/P_h}{\sqrt{\sum N_h^2/P_h}} \end{aligned} \tag{D.16}$$

Again, q is determined from the total cost constraint. Now consider a given program, and let

$$P_h = (1 + f_h)P_h^* \tag{D.17}$$

where P_h denotes the compromise selection probability for area h and P_h^* the optimum selection probability for the given program. Let n_h^* denote the corresponding numbers of enrollees to be taken for the sample. Then

$$\frac{n_h}{n_h^*} = \frac{1}{1+f_h} \qquad \frac{\sqrt{\Sigma \frac{N_h^2}{P_h^*}}}{\sqrt{\Sigma \frac{N_h^2}{(1+f_h)P_h^*}}} . \tag{D.18}$$

D.6 According to (D.14), with the optimum selection probabilities, a constant number of enrollees would be sampled per area. The implication of (D.19) is that, if the size of program in an area is relatively small compared to that implied by the selection probability used (i.e., NP_h), the number of enrollees to be sampled should be smaller than the average per area; if the size of program is relatively large, the number of enrollees to be sampled should be larger than the average per area. Under conditions for which sampling with probability proportional to program size would be optimum, the reduction or increase would be in direct proportion to the ratio of the program size to that implied by the selection probability, maintaining the self-weighting condition; that is

$$\frac{n_h}{N_h^*} = \frac{N_h}{NP_h} . \tag{D.19}$$

D.7 From the point of view of analyses for the individual areas, it would be desirable to equalize the enrollee sample by area and by program within area (assuming equal variance components). Nevertheless, in actual fact, it is to be expected that in particular programs in the individual study areas the flow of enrollees during the study sampling period will be too small to meet this objective; and the deficit will be balanced over the total sample by sampling larger than average numbers of enrollees from areas with relatively large programs. Also, the measures of program size were, of necessity, based on data for the fiscal year preceding that in which the study sample will be selected. The implications of this analysis are that:

a. These adjustments should be distributed over the study areas in accordance with expected program size to the extent feasible

b. A fortiori, adjustment of the enrollee sampling by program and area in this way should tend to improve the statistical efficiency of the study design, given area selection probabilities actually used, for national projections from the study and for projections to the study areas.

Optimum Probabilities Compared with Stratification of Areas by Size

D.8 The same approach is applicable to a sampling model in which areas are assumed to be stratified by size and sampled with equal probabilities within stratum.[2] In that case, the optimization determines the sampling rate for areas to be used in each of the size strata. The optimum sampling rates for areas in the different size strata mimic the optimum selection probabilities for areas in the unstratified model so that the probability of selection for a given area is essentially the same function of size under either sampling model. This is the basis for the statement in paragraph 3.20 that the two approaches yield the same effect.

Implications of Alternative Objectives

D.9 The objective in the analysis approach followed, i.e., with the number of areas to be selected not fixed, was to permit some exploration of the optimum number of areas assuming no constraint in the initial specifications for the study other than total cost. It may be remarked that, on the available evidence, if the primary objective of the study were national projections, the optimum number of areas to sample should be greater than ten.

SPECULATED SAMPLING ERRORS

D.10 The following details supplement paragraphs 3.23 - 3.40.

The Speculated Sampling Errors

D.11 For the sampling error investigation the 10-area design outlined in paragraph 3.21 was assumed, with the statistic of interest being average annual post-program earnings per enrollee. Consider a given program, say program g. The estimator used for the investigation is:

$$\bar{\bar{x}}_{rg} = \frac{1}{N_g} \quad N_{g1}\bar{\bar{x}}_{rg1} + N_{g2}\bar{\bar{x}}_{rg2} + N_{gs}\bar{\bar{x}}_{rgs} \qquad \text{(D.20)}$$

where

$$\bar{\bar{x}}_{rgs} = \frac{\sum_{h=3}^{10}\sum_{i=1}^{1} \frac{N_{ghi}}{P_{hi}} \bar{\bar{x}}_{rghi}}{\sum_{h=3}^{10}\sum_{i=1}^{1} \frac{N_{ghi}}{P_{hi}}} . \qquad \text{(D.21)}$$

$\bar{\bar{x}}_{rg1}$ and $\bar{\bar{x}}_{rg2}$ are estimates of average annual post-program earnings per program g enrollee in each of the two self-representing strata, and $\bar{\bar{x}}_{rgs}$ is a combined ratio estimate of the average in the eight nonself-representing strata; and

---

[2] See the references cited in Footnote 1.

N_{g1}, N_{g2}, and N_{gs} are corresponding assumed populations to which the estimates are to be projected, with

$$N_g = N_{g1} + N_{g2} + N_{gs} \,. \tag{D.22}$$

There are alternative estimators which might be useful, and these will be considered in the study. There are also a number of alternatives for constructing the averages $\bar{\bar{x}}_{ghi}$. For this analysis it is assumed that a post-program earnings figure for the time period of interest is available for each sample individual, from which the average per enrollee is computed. So

$$\bar{\bar{x}}_{ghi} = \frac{x_{ghi}}{n_{ghi}} = \frac{1}{n_{ghi}} \sum_{j=1}^{n_{ghi}} x_{ghij} \tag{D.23}$$

where there are n_{ghi} sample individuals from program g in area i of strata h. The sampling variance of $\bar{\bar{x}}_{rg}$ is found by standard methods to be

$$\sigma^2_{\bar{\bar{x}}_{rg}} = \Pi_g^2 \frac{1}{N_{gs}^2} \sum_{h=3}^{10} \sum_{i=1}^{Q_h} P_{hi} Z^2_{ghi} + \frac{1}{N_g^2} \sum_{h=1}^{10} \sum_{i=1}^{Q_h} \frac{1}{P_{hi}} \frac{N^2_{ghi}}{n_{ghi}} \sigma^2_{ghi} \tag{D.24}$$

where $\Pi_g = N_{gs}/N_g$

$$Z_{ghi} = \left\{ \left(\frac{N_{ghi} \bar{\bar{x}}_{ghi}}{P_{hi}} - N_{gh} \bar{\bar{X}}_{gh} \right) - \bar{\bar{X}}_{gs} \left(\frac{N_{ghi}}{P_{hi}} - N_{gh} \right) \right\}$$

with $\bar{\bar{X}}_{ghi} = E\left(\bar{\bar{x}}_{ghi}/ghi\right) \qquad \bar{\bar{X}}_{gh} = \sum_{i=1}^{Q_h} N_{ghi} \bar{\bar{X}}_{ghi}/N_{gh}$

$$\bar{\bar{X}}_{gs} = \sum_{h=3}^{10} N_{gh} \bar{\bar{X}}_{gh}/N_{gs}$$

$$N_{gh} = \sum_{i=1}^{Q_h} N_{ghi} \qquad N_{gs} = \sum_{h=3}^{10} N_{gh}$$

and $\sigma^2_{ghi} = E\left\{\left(x_{ghij} - \bar{\bar{X}}_{ghi}\right)^2 /ghi\right\}$.

Q_h denotes the number of areas in stratum h, and N_{ghi} an assumed population in area (hi) to which the estimate for program g is to be projected. For the analysis it was assumed that

$$\Pi_g = 0.8. \tag{D.25}$$

Then, with a self-weighting sample

$$\sigma^2_{\bar{\bar{x}}_{rg}} = (0.8)^2 \frac{\sigma^2_{Bg}}{8} + \frac{\sigma^2_{Wg}}{n_g} \tag{D.26}$$

where

$$\sigma^2_{Bg} = \sum_{h=3}^{10} \sum_{i=1}^{Q_h} P_{hi} Z^2_{ghi} / 8 \bar{N}^2_{gs} \tag{D.27}$$

$$\sigma^2_{Wg} = \sum_{h=1}^{10} \sum_{i=1}^{Q_h} N_{ghi} \sigma^2_{ghi} / N_g \tag{D.28}$$

with $\bar{N}_{gs} = N_{gs}/8$, and n_g the total size of sample from program g. Equation (D.26) was used for computing the speculated sampling errors for average annual post-program earnings per enrollee in a given program presented in the text tables. Numerical values to use for σ^2_{Bg} and σ^2_{Wg} in (D.26) were taken from generalized variance curves, the derivation of which is described in Appendix E. For the speculated sampling errors of average annual post-program change in income compared with controls, the change is $(\bar{\bar{x}}_{rg} - \bar{\bar{x}}_{rc})$ where $\bar{\bar{x}}_{rc}$ is the average annual income per control case comparable to post-program income of enrollees, with sampling variance

$$\sigma^2_{\bar{\bar{x}}_{rg} - \bar{\bar{x}}_{rc}} = \sigma^2_{\bar{\bar{x}}_{rg}} + \sigma^2_{\bar{\bar{x}}_{rc}} - 2\rho_{\bar{\bar{x}}_{rg}\bar{\bar{x}}_{rc}} \sigma_{\bar{\bar{x}}_{rg}} \sigma_{\bar{\bar{x}}_{rc}} \tag{D.29}$$

where $\rho_{\bar{\bar{x}}_{rg}\bar{\bar{x}}_{rc}}$ is the correlation between the estimates of annual post-program earnings of enrollees and controls. Each of $\sigma^2_{\bar{\bar{x}}_{rg}}$ and $\sigma^2_{\bar{\bar{x}}_{rc}}$ has the form (D.26) so that, if it is assumed that annual earnings of enrollees and controls are correlated only at the area level,

$$\left.\begin{aligned}\sigma^2_{\bar{\bar{x}}_{rg}-\bar{\bar{x}}_{rc}} &= \frac{(0.8)^2}{8}\left\{\sigma^2_{Bg}+\sigma^2_{Bc}-2\rho_{Bgc}\,\sigma_{Bg}\,\sigma_{Bc}\right\}\\ &+\left\{\frac{\sigma^2_{Wg}}{n_g}+\frac{\sigma^2_{Wc}}{n_c}\right\}\end{aligned}\right\} \qquad (D.30)$$

where σ^2_{Bc} σ^2_{Wc} are defined as in (D.27) and (D.28) for the control group.

For deriving numerical estimates, the first term of (D.30) was approximated by

$$\frac{(0.8)^2}{8}\left\{\sigma^2_{Bg}+\sigma^2_{Bc}\right\}(1-\rho). \qquad (D.31)$$

Again, numerical values to use for σ^2_{Bc} and σ^2_{Wc} were taken from generalized variance curves, the derivation of which is described in Appendix E.

Application of Speculated Sampling Errors

D.12 The interpretation of the sampling errors in paragraph 3.35, and in the form illustrated in paragraphs 3.40 and 7.9, i.e.,

> The chances are about 2 out of 3 that the difference between the (sample estimate) and the (statistic) which would be found from a study of all enrollees would be less than --(sampling error)

and

> The chances are about 19 in 20 that the difference between the (sample estimate) and the (statistic) which would be found from a study of all enrollees would be less than twice -- (sampling error)

is a standard form of confidence interval statement. The "chances" cited assume that the sample estimate referred to is approximately normally distributed. It is anticipated that the study sample will be sufficiently large that this assumption will be tenable for the major independent estimates.[3] The "study of all enrollees" referred to assumes the use of the same data collection instruments and procedures as for the sample covered in the study that is carried out, the only difference between the two being in the "sample" size. For this reason, the "study of all enrollees" is sometimes referred to as "equal complete coverage."[4/]

---

[3] See P. Erdos and A. R'enyi, "On the Central Limit Theorem for Samples From a Finite Population," Pub. Math. Inst. Hungarian Acad. Sci., 4(1969), 49-57 and J. H'ajak, "Limiting Distributions in Simple Random Sampling From a Finite Population," Pub. Math Inst. Hungarian Acad. Sci., 5(1960), 361-374 for theoretical conditions; and W.G. Cochran, Sampling Techniques, 2nd Ed., 38-44, John Wiley & Sons, New York, 1963 for a discussion of the validity of the normal approximation.

D.13 The following details supplement the discussion of the effects of measurement error in paragraph 3.35. Since the sampling error does not include the potential contribution of systematic errors arising from the design of the data collection instruments and procedures or their execution, it may not provide an adequate measure of uncertainty if the study estimates are subject to bias. The mean square error (MSE) of an estimate, measured from the population value that is to be estimated, does include the contribution of such errors. It may be expressed as

$$MSE = \sigma^2 + B^2 \tag{D.32}$$

where σ = sampling error of the estimate

and B = bias of the estimate.

A rule of thumb frequently recommended is that if $B/\sigma < 0.1$, confidence interval statements for a normally distributed but biased estimate based on σ will have approximately the same confidence (or odds) as if based on $\sqrt{MSE}$.[5/] As the bias increases relative to sampling error, probabilities of errors may be rapidly affected, particularly the probabilities for one-sided (i.e., positive or negative) change which are computed from only one tail of the distribution of the estimate. If B/σ is less than about 0.5 confidence interval statements for a biased estimate based on $\sqrt{MSE}$ will have approximately the same confidence (or odds) as corresponding statements for an unbiased estimate based on σ.[6/]

D.14 What is called the OC (Operating Characteristic) curve of the study in paragraph 3.37 gives the probability of finding a statistically significant post-program increase in earnings of enrollees compared with controls based on the study sample when there in fact is a "true" post-program increase of the magnitude specified. By "true" increase is meant the increase that would be found from a study covering all program enrollees. As noted in paragraph 3.37 the statistical test is that of the hypothesis that there is no increase in earnings (H_0), against the alternative that there is an increase. The curves are calibrated so that the probability of finding a statistically significant increase based on the study sample when there in fact is no true increase is 0.05. What is called the OC curve here is usually referred to as the power

---

[4/] See, for example, W.E. Dening, Sample Design in Business Research, John Wiley & Sons, Inc., New York, 1966, Chapter 1.

[5/] Compare, for example, W.G. Cochran, op. cit., Chapter 1.

[6/] Hansen, Hurwitz and Madow, op. cit., Vol I, Chapter 2.

function of the statistical test involved; the OC curve cited here is the complement of the power function in this terminology. The OC curve computations given in this report assume that the study estimates are approximately normally distributed.

Implications for Benefit-Cost Ratios

D.15 The relationship used to arrive at speculated sampling errors for benefit-cost ratios is derived as follows . Since $r = B'/C'$ we can write

$$B' = rC' \tag{D.33}$$

so that

$$V^2_{B'} = V^2_{rC'} \tag{D.34}$$

since rC' is the product of two random variables, $V^2_{rC'}$ is [7]

$$V^2_{rC'} \doteq V^2_r + V^2_{C'} + 2\rho_{rC'}V_rV_{C'} \tag{D.35}$$

and by the assumption that the benefit-cost ratio and program cost are uncorrelated

$$\rho_{rC'} = 0$$

substituting (D.35) in (D.34) with this assumption

$$V^2_{B'} \doteq V^2_r + V^2_{C'} \tag{D.36}$$

so that

$$V^2_{B'/C'} \doteq V^2_{B'} - V^2_{C'} \tag{D.37}$$

Sample Allocation With a Fixed Total Cost

D.16 For approximating an optimum allocation between the program enrollee and control group samples, we note first that for purposes of national projection of the study results, control cases should be sampled in each area in which program enrollees are sampled. This provision is dictated not only by considerations of sampling error, but by the fact that it is likely that area-environment influences on the outcomes observed for program enrollees cannot be adequately quantified. Thus, matching within area--essentially a nonparametric adjustment technique for the area-environment factors--is desirable. With this stipulation, the optimization can affect only the within-area component of the variance of national projections. In the case of analyses for the study areas as a fixed set, of course, this is the only component of variance.

---

[7] For a general discussion of the variances of products, see Leo A. Goodman, "The Variance of the Product of k Random Variables," J.Amer. Statist. Assn., 57(1962), 54-60.

D.17 From equation (D.30), the within area component of the variance of the estimated average annual post-program change in earnings for a given program is

$$\frac{\sigma^2_{Wg}}{n_g} + \frac{\sigma^2_{Wc}}{n_c} \qquad \text{(D.38)}$$

and the within area cost is assumed to be represented by the cost function

$$c_g n_g + c_c n_c \qquad \text{(D.39)}$$

where c_g and c_c are the variable costs per program enrollee and per control case in sample, respectively. To derive the optimum allocation, form the Langrangian

$$\psi = \frac{\sigma^2_{Wg}}{n_g} + \frac{\sigma^2_{Wc}}{n_c} + \lambda\ (c_g n_g + c_c n_c - C) \qquad \text{(D.40)}$$

where C is the assumed fixed total cost. Proceeding as in paragraph D.3, the optimum allocation is found to be

$$\frac{n_c}{n_g} = \sqrt{\frac{\sigma^2_{Wc}/C_c}{\sigma^2_{Wg}/C_g}} = \sqrt{\frac{\sigma^2_{Wc}}{\sigma^2_{Wg}} \Big/ \frac{C_c}{C_g}}\ . \qquad \text{(D.41)}$$

Distinctions were not made between the optimum allocations with regard to the different programs because it was concluded that there was not sufficient information to warrant firm distinctions between the values of $\sigma_{Wg}\sqrt{c_g}$ for the different programs.

D.18 From the data presented in Appendix E, the ratio $\sigma^2_{Wc}/\sigma^2_{Wg}$ was taken to be approximately 1.4. To evaluate the unit costs, C_g and C_c, it was necessary to make some assumptions as to the structure of the unit costs for program enrollee cases and control cases. The assumptions as to loss rates are summarized in paragraph 3.44, and the cost parameters in paragraph E.18, Appendix E. With these assumptions, the total field costs to obtain data at the end of the study for one control case would be on the order of 2.4-2.8 those for one enrollee. Since a single control case will serve for all programs, for each of the two age groups, this cost ratio should be divided by the number of programs served to obtain a cost ratio attributable to a particular program. For the age 22 and over group, a control case will serve for both the JOBS and MDTA (Inst) programs. This leads to the conclusion that n_c/n_g should be about 1.0-1.1, or that the sample size for each of the programs and for the control group after attrition should be about the same. For the under 22 age group, the conclusion would be that n_c/n_g should be about 1.2-1.4. Considering that the optimum is likely to be fairly broad, and that the program cases will be the base for other analyses of interest, it is recommended that the sample size for each of the programs and for the control group after attrition should be the same for the under 22 age group also.

SAMPLING PROGRAM ENROLLEES AND CONTROL CASES

D.19 The following details supplement Section V.

Sampling of Program Enrollees

D.20 It should be noted that enrollees will be selected by a two-stage procedure rather than by simple random sampling as assumed in the computation of the speculated sampling errors. The effect of this for ratios or averages, in relvariance terms approximately is to multiply the within-area component of the relvariance that would have been obtained with simple random sampling by the factor

$$[1 + \delta (\bar{n} - 1)] \tag{D.42}$$

where $\bar{n}$ is the average number of enrollees sampled per program project and δ is the measure of homogeneity for enrollees within project. A target $\bar{n}$ on the order of 8 will be aimed for (see Appendix E). The ability to achieve this will depend on the numbers of projects enrolling individuals during the study sampling period. However, the two-stage procedure is not expected to lead to relative increases of more than a few percent in the sampling variances of study estimates since the effect of the proposed stratification at the first stage will be to reduce δ, unless the entire sample must be selected from only one or two projects. The impact on totals estimated from the study, if any, would be somewhat greater.

D.21 The decision to not employ controls on the numbers of individuals by sex and race selected from an individual projects, or to impose other controls in the sampling of enrollees, is based on the desire to keep the within-area analysis within the limits of available statistical tools. The use of estimation techniques based on post-stratification for estimating post-program changes in earnings will be considered in the study if it appears desirable.

Sampling of Controls

D.22 Scheduling of Screening. When the screening for the control sample is not carried out simultaneously with the sampling of program enrollees, there is a potential for bias which is not discussed in paragraph 5.13. If screening is done in advance of the enrollee sampling, the potential for bias appears likely to favor the control group. To see this, suppose the screening were done at time t_1, and selection of enrollees at time t_2 as they enroll. An individual in the population who was unemployed at time t_1 could be selected as a control case and would then be retained in the sample even if he were employed at time t_2. On the other hand, it would seem that at least some individuals unemployed at time t_1 but employed at time t_2 would be eliminated as potential program enrollees. If the screening follows the enrollee sampling, the potential bias appears likely to favor the enrollee group. If the time lag between screening and enrollee sampling were not long, the impact of this potential for bias would undoubtedly be small.

D.23 Matching in Selection of Controls. The technique of matching controls with enrollees on the basis of a number of variables is a nonparametric technique for reducing the variance of enrollee-control comparisons and, in particular, of estimated post-program changes in earnings. It is nonparametric in the sense that it does not depend on the mathematical from of the relationship between the variables used for matching and the variable of interest. The efficiency of matching depends upon the (multiple) correlation between the matching variable(s) and the variable of interest. Thus, it is tempting to select controls with matching on a fairly extensive list of characteristics.

D.24 To attempt matching on a relatively large number of variables would be time-consuming and costly for the screening operation required to obtain "acceptable" control individuals. Further, it is likely to result in fairly distorted control samples.

D.25 If the program enrollees reflect selection in the recruiting of individuals for the program, the estimates of program impacts based on matched controls are likely to give a biased estimate of the program effects to be expected for the target population.

D.26 As shown in Appendix F, regression analysis can be expected to provide about the same precision (sampling variance) as matching with the sample sizes contemplated. This assumes that the correct regression function is used; otherwise, a bias term must be added. If the correct regression function is used, the fact that program enrollees represent a restricted selection from the target does not introduce a bias in the estimation of program impacts, but only an increase in variance.

D.27 Enrollees and controls will be matched on race, sex, and broad age group. Since both groups will meet program eligibility requirements, it is expected that the range of variables of interest for regression analysis will generally be constrained in comparison with the population at large. Therefore, it is expected that linear regressions (or simple modifications of them) will generally provide at least a good approximation model.

D.28 The consideration being given to matching by geographic area within the LMA is motivated by the fact that it is not adequately known how the associated area-environment can be quantified for a parametric (i.e., regression) approach.

D.29 Central City-Balance of SMSA Distribution. As background for the discussion in paragraph 5.19, the distribution of the poverty population in 1966 by race and location within SMSAs was estimated to be as follows:

Race	Percent Distribution by Location		
	Total	Central City	Balance of SMSA
White	67	57	83
Non-white	33	43	17
Total	100	100	100

Source: Special tabulation of the March 1967 CPS for OEO.

Further data are provided by the Sample Urban Employment Surveys (UES) covering the CEP areas of Atlanta, Chicago, Detroit, Houston, Los Angeles, and New York City, plus additional target neighborhoods in New York. For the combined UES areas, averaged over the quarter of 1968, the following racial distribution was estimated.

Race	Percent Distribution of Persons 20 and Older by Status	
	Civilian Non-institutional Population	Unemployed
White	33	23
Non-white	67	77
Total	100	100

Source: BLS News Release USDL-10-278, February 20, 1969.

D.30 Target Population Sampling. The discussion in paragraphs 5.21-5.22 of potential undercoverage of the target population in household screening draws upon data from the Pilot and Experimental Program on Urban Employment Surveys sponsored by the Manpower Administration,[8] and from a test conducted in connection with a special census of Trenton, New Jersey in 1968 by the Bureau of the Census.[9] In both experiences, lists were created by contacting men in places they are thought to frequent (e.g., bar, restaurant, poolroom, barber shop, street corner) and obtaining their names and addresses. In the UES studies, interviews were then conducted at the given addresses, as found possible, to determine whether the men would be reported in a household survey. In the Census test, the census listing were checked to see whether the men had been enumerated. The data reported cannot be summarized simply.

---

[8] See, Bureau of Labor Statistics, Pilot and Experimental Program on Urban Employment Surveys, Report No. 354, March 1969.

[9] See, Bureau of the Census, Response Research Branch, Missed Persons Campaign in Trenton, New Jersey, Report No. 69-8, 26 February 1969 (Revised).

APPENDIX E

NUMERICAL ESTIMATES OF VARIANCE AND COST PARAMETERS

SOURCES OF DATA

Variance Parameters

E.1 The data most directly relevant to developing numerical estimates of variance parameters for design of the study sample are previous data on annual pre-program and post-program earnings, on an individual person basis. It is exceedingly fortunate that OEO was able to arrange access for the study to such data from a national evaluation study of Job Corps terminees conducted by Louis Harris and Associates, Inc., in 1969, and DOL to data from a national study of NYC Out-of-School terminees conducted by Dunlap and Associates, Inc., in 1967. The data in both studies are based on recall by the respondent. These studies were the primary source of data for estimates of variance parameters. Only limited use of program statistics was possible because, by and large, they do not provide longitudinal data.

Cost Parameters

E.2 Numerical values for cost parameters are based on cost estimates provided by the National Opinion Research Center of the University of Chicago (NORC), the survey contractor for the study.

VARIANCE PARAMETERS

E.3 For design of the enrollee sample, estimates were wanted for the levels of variance to be expected between enrollees within program project, between projects within area, and between areas. The estimates which were developed are described in the following paragraphs.

Variance Between Enrollees Within Program Project

E.4 Variance Estimates. Estimates of the relvariance in annual post-program earnings between NYC Out-of-School program terminees within site (project) were computed using the Dunlap survey data from the following 28 urban sites.[1]

Sites in SMSAs Comparable in Size to the Study SMSAs	Sites in Smaller SMSAs
New Haven, Conn.	New Bedford, Mass.
Bridgeport, Conn.	New Britain, Conn.
Patterson, N.J.	Pawtucket, R.I.
Troy, N.Y.	Lynn, Mass.
White Plains, N.Y.	Pine Bluff, Ark.
Louisville, Ky.	Waco, Tex.
Philadelphia, Pa. (2 sites)	Baton Rouge, La.
Chattanooga, Tenn.	Little Rock, Ark.
Miami, Fla.	Cedar Rapids, Iowa
Minneapolis, Minn.	Cheyenne, Wyo.
Grand Rapids, Mich.	Brighton, Colo.
Milwaukee, Wis.	St. Joseph, Mo.
Oakland, Calif.	
San Jose, Calif.	
Spokane, Wash.	

The characteristics of the terminee sample used for the computations are shown in Table E.1, and the estimates of relvariance in Table E.4 . The individual post-program earnings of the terminees were computed by LS&R from data in the survey giving, for each period of employment, the number of weeks employed, the average hours worked per week, and the average hourly wage rate paid. Time spent in other manpower programs, military service, or school, is not included in the computation of earnings. The effect is to underestimate average earnings. Because the survey covered varying lengths of post-program experience, the estimates of relvariance correspond to a ratio estimate of annual earnings based on earnings per week covered in the survey. The average number of weeks reported for, per terminee, is approximately 42. Relvariances were computed from the usual approximation of the form

$$\nu^2_{\frac{x}{y}} = \nu^2_x + \nu^2_y - 2\,\nu_{xy}$$

---

[1] The Dunlap study also included a sample of "controls." This sample, however, was not large enough for the computations here. Data from three urban sites were not included because the sites could not be used also for the estimates of relvariance between sites that were computed.

where

$$\nu_x^2 = s_x^2 / x^2 \qquad \nu_{xy} = s_{xy} / xy \qquad \nu_y^2 = s_y^2 / y^2$$

The relvariances in Table E.4 are pooled estimates. The estimated variances and covariances are pooled averages of the individual within-site estimates. The same is true for the denominators of the relvariances. Separate estimates, not shown in Table E.4, were also computed for sites in the larger SMSAs and the smaller SMSAs for more detailed analysis.

E.5 Estimates of the relvariance in annual pre-program and in post-program earnings between Job Corps terminees, within groups defined by Census geographic region and 1960 population of place in which the terminee was found for interview, were computed using the Harris survey data. These estimates presumably include some between-place component and are therefore overestimates to an unknown extent of within-place variances. However, since in the study Job Corps enrollees need not return to the city from which they entered the program, the study estimates will also be subject to this component. The characteristics of the terminee sample used for the computations are shown in Tables E.2 and E.3, and the estimates of relvariance in Table E.5. The individual post-program annual earnings data by terminee are estimates by Harris based on terminee reports covering their first six post-program months of experience. The relvariances for each of pre-program and post-program earnings are based on the reports for each, regardless of whether both figures were reported. Blank entries were treated as "not reported."

E.6 The data summarized in Tables E.1 through E.5 showed several patterns which suggested an approach to "smoothing" the variance information from the two programs and generalizing it to all five programs. First, it should be noted that the variance estimates are themselves subject to sampling variance, and some "smoothing" would be reasonable. Also, generalizing the variance estimates might provide some insight into what would happen to the study data under alternative economic developments during the 18-month follow-up period.

E.7 The patterns in Tables E.1 through E.5 of interest are, first, a tendency for the relvariances to be smaller for groups with higher average earnings. This was investigated with the NYC Out-of-School terminee data. The results, shown in Table E.6, indicate, as expected, that the major component of the total variance of average earnings between trainees arises from variation in patterns of working rather than from earnings per week when employed.[4] Second, the Job Corps data show higher earnings, both pre-program and post-program than the NYC Out-of-School data. Thus, they might be useful for scaling up the variance estimates from the two programs to higher levels that might be observed with the JOBS and MDTA programs.

---

[4] This would not necessarily be expected, of course, for groups of individuals more regularly employed on a full-time basis.

E.8 Variance Curves. The approach followed to generalize the variances estimates assumes that as earnings per week increase, the pattern of working will stablize in accordance with the patterns found in the two programs analyzed. The Job Corps data available to LS&R did not show number of weeks worked, but permitted distinguishing only those individuals who did not work at all from those who had any employment. Accordingly, curves were developed by eye-fit to the data for the proportion of terminees expected to have any employment as a function of their average earnings per week; and for the relvariance in earnings per week among those having any employment, again as a function of their average earnings per week. The estimates of relvariance by race and sex were treated as individual observations without regard to race and sex for this purpose. There was the possibility, therefore, that the variance behavior observed at different levels of average weekly earnings might be an artifact of characteristic differences between the race-sex groups. Some check on this possibility was made by comparing the points for each race-sex group for each of the pre-program and post-program earnings of Job Corps terminees that were obtained from the separate relvariance estimates for the larger cities and for the smaller cities. The conclusion was that speculating as to sampling variances on the basis of earnings would be better than on the basis of race-sex groups, even though average earnings in the data analyzed were related to race and sex. Tying the generalized variance curves to annual earnings rather than the characteristics of the individual person could have an advantage if the data from the surveys used were affected by nonresponse biases. For example, the surveys were retrospective, so that it may be speculated that it would be easier to locate and obtain data from more successful individuals than from less successful individuals.[3/] In this event, the patterns of employment of those individuals for whom data were obtained might be more representative of individuals having similar annual earnings than of individuals having the same demographic characteristics such as race and sex. The two curves were then combined to develop curves for the expected relvariances between terminees within program project as a function of average weekly earnings.

E.9 The NYC Out-of-School terminee data and the post-program Job Corps data both represent program enrollees. The pre-program Job Corps data were used as a guide to what the variance characteristics of control cases in the study might be. The generalized relvariances for average pre-program earnings were approximately 50 percent higher than for corresponding average post-program earnings. Some limited comparisons between the controls and the terminees with controls in the NYC Out-of-School data indicated relvariances on the order of 40 percent higher for the controls than for the terminees even though average annual earnings were about the same. It is possible to think of explanations for this phenomenon, but the reason why this should be so is not known.

[3/] This is certainly not always the case, particularly in the case of the Job Corps terminees, since "success" of an individual may be associated with his moving which would make it harder to find him.

E.10 Table E.7 provides a summary comparison between the relvariances computed from the available data and the corresponding figures from the generalized variance curves. The table shows the percent differences in the coefficients of variation, and thus the percent differences in the sampling errors of estimates that would result from samples based on the two figures. A dash indicates a difference of less than one percent. For example, an entry of 10 (or -10) in the table indicates that the sampling error that would be predicted from the generalized curves would be 10 percent higher (or 10 percent lower) than that predicted from direct estimates from the survey data analyzed. If the predicted sampling error based on the direct estimates from the survey data analyzed were, say, $100,the prediction with 10 percent (or -10 percent) error would be $110 (or $90) The agreement between the two sets of figures is considered satisfactory. The conclusion is that the variance curves should be considered as only approximations, but, that they provide a reasonable fix on the general within-project sampling behavior to be expected in the study estimates.

E.11 The variance curves developed for the program enrollees and controls are graphed in Figure E.1.[4/]

Variance Between Projects/Areas

E.12 The data available for developing similar speculations for the between project and/or area component of the total sampling error to be expected in making national projections were more limited.

E.13 The principal data available were from the NYC Out-of-School terminee survey. Sites from which data were used for estimating within-site relvariances were stratified by size of place and geographic region in which located. This led to six strata for the larger urban sites and three strata for the smaller ones. Estimates were made of average between and within site components of variance, within stratum, for the statistic average earnings per week. The between component estimated corresponds to a stratified sample of sites with probability of selection porportional to program size (or a measure highly correlated with it). A number of the between component estimates were negative. These were treated as zeros in the averaging.[5/] The resulting estimates of average between-components are summarized in Table E.8. Additional estimates were made from the Job Corps terminee survey data. These were more questionable with regard to levels of variance, since only size of place in which the terminee was interviewed was identified rather than specific place. Rather, they were used to check the relationships between the levels of variance for the race-sex groups.

[4/] Presumably, if extended to higher annual earnings, the two curves should coincide with each other and with a similar curve which might be derived for individuals with earnings levels above those of the program target populations.

[5/] This is a biased but minimum mean square procedure. See M.G. Kendall and Alan Stuart, The Advanced Theory of Statistics, Hafner Publishing Co., Vol.3., 1966, p. 71. B.J. Tepping, " Note on Restricted Estimates," U.S. Census, Technical Notes, No. 1, 31-33, Washington, D.C.,1968.

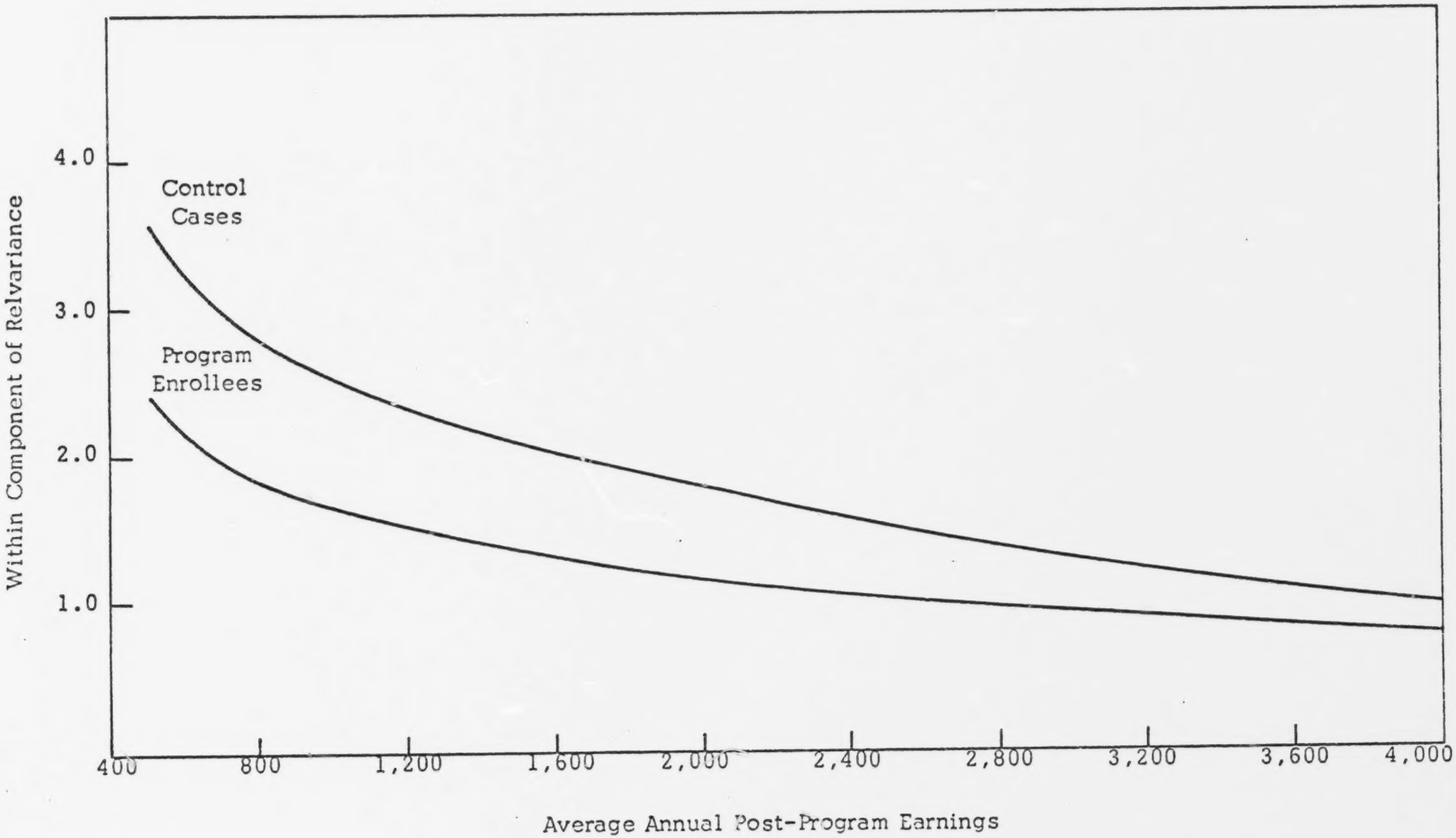

FIGURE E.1. GENERALIZED WITHIN COMPONENT OF RELVARIANCE PER PERSON, AVERAGE ANNUAL POST-PROGRAM EARNINGS: TERMINEES AND CONTROLS

Finally, a 10-stratum sample of the study universe SMSAs was established; and a between SMSA within stratum variance computed for 1967 average weekly earnings of production or nonsupervisory employees in manufacturing industries as reported in Employment and Earnings Statistics for States and Areas, 1939,67, Bulletin No. 1370-5, Bureau of Labor Statistics, August 1968. The purpose of this computation was to help scale the survey estimates up to higher earnings levels. The generalized variance curve derived for the between-component is plotted against that for the within-component in Figure E. 2 . The between-component is assumed to be the same for both program enrollees and control cases.

E.14 Variance Relationships. The variance relationships shown by Figure E.2. are illustrated by the following examples of the estimated fraction of the total relvariance (sum of the two components) due to the between component:

Type of Case	Level of Annual Earnings	
	$1,000	$4,000
Program enrollees (post-program)	.025	.017
Control cases	.017	.013

This fraction may be roughly interpreted as the intraclass correlation or measure of homogeneity, say δ, of earnings within projects. The quantity $\sqrt{(1-\delta)/\delta}$ is a basic parameter for speculating the number of cases per area for an optimum study design. Smaller values of this parameter indicate smaller numbers of cases per area as optimum. If all projects in an area are sampled and the between-component is interpreted as a between-area component, the values of this parameter corresponding to the fractions in the tabulation above are as follows:

Type of Case	Level of Annual Earnings	
	$1,000	$4,000
Program enrollees (post-program)	6.2	7.6
Control cases	7.6	8.6

E.15 To provide some further background, estimates of the fraction of the total relvariance due to the between area component were made for the

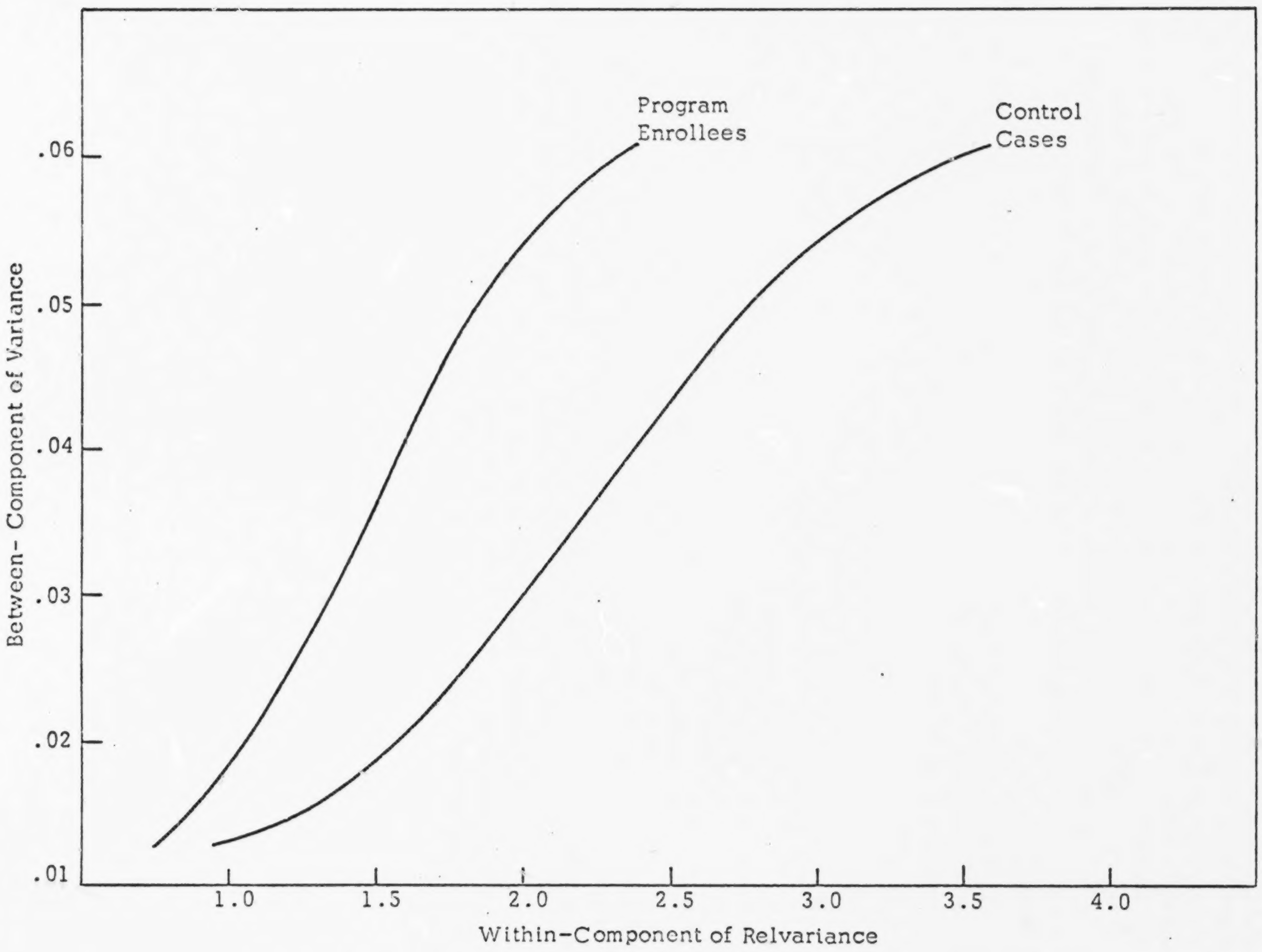

FIGURE E.2. RELATIONSHIPS OF GENERALIZED BETWEEN AND WITHIN COMPONENTS OF VARIANCE, AVERAGE ANNUAL EARNINGS

dropout rate in MDTA (Institutional Component) in FY 68 and in JOBS (Contract component) in the SMSAs in the study universe. These data were obtained from special tabulations made for the study by the Office of Manpower Data Systems, DOL. The computations were carried out for a 10-stratum and a 15-stratum design, and assume simple random sampling within area. The values of $\sqrt{(1-\delta)/\delta}$ obtained are as follows: [7]

Type of Design	Program	
	MDTA (Inst.)	JOBS (Contract)
10-stratum	5.4	5.0
15-stratum	6.7	6.1

The conclusion is that the minimum of the variance function for estimates of items from the study such as those analyzed is likely to be fairly broad in the neighborhood of the optimum number of cases per area. Thus, a compromise number would be approximately optimum for each of them.

E.16 Correlation Relationships. For estimates of differences in earnings between program enrollees and controls, and between program enrollees in alternative programs in which they might have been enrolled, the correlation of earnings between the groups compared is a parameter of the sampling error of the difference. It is assumed that there is no correlation between the groups within area for a fixed time period. Therefore, the correlation of interest is that between groups at the area level. Using the NYC Out-of-School survey data, an estimate was made of the correlation at the site level (within strata) between the controls and terminees with controls for average earnings per week. For sites in the larger urban places the estimated correlation was 0.65, and for those in the smaller urban places, 0.51. For purposes of speculation, it should be noted that the average weekly earnings of the controls and the terminees with controls in the data used were practically identical. Thus, the correlation estimates may be high. Perhaps a maximum value of 0.5 might be used for purposes of speculation for program-control comparisons, and a value of 0.3 might represent a more conservative expectation for both program-control and program-program comparisons.

COST DATA

E.17 The cost data cited here are estimates provided by NORC as guides to cost relationships for developing the study sample design, and are based on only preliminary specifications for the field work. They cannot be used to construct estimates of the study cost because of the omission of components of cost not entering into sample design trade-offs.

---

[7] The data indicated substantial variability in dropout rates between MDTA classes within area, and between JOBS contracts within area. This would suggest that if projects are sampled, larger numbers of cases per area than indicated by the values in this table would be optimum.

E. 18 The parameter for overhead cost per area is $36,000. The parameters for interviewing costs are as follows:

Survey Operation	Unit Cost Per Person in Initial Sample	
	Program Enrollees	Control Cases
Sampling	\$ 2	\$ 7 1/
Pre-program interview	\$ 5 2/	\$15
Post-program interview	\$ 5 2/	\$ 7
Follow-up interview	\$ 6 3/	\$ 7 3/
Total	\$30	\$50

1/ Includes listing and screening.
2/ Assumes major proportion of interviews completed at project sites.
3/ Cost per person in initial sample for each of 3 follow-up interviews. Cost per person completed in successive follow-up interviews increases.

TABLE E.1

CHARACTERISTICS OF TERMINEES FROM NYC OUT-OF-SCHOOL PROGRAMS IN SITES USED FOR ANALYSIS[1,2]

Race of Terminee	Sex of Terminee		
	Male	Female	Total
1. Average Earnings Per Week			
White	$28.96	$17.52	$23.97
Non-white	$21.65	$16.39	$18.48
Total	$24.83	$16.72	$20.43
2. Average Earnings Per Week Worked			
White	$54.89	$51.09	$53.64
Non-white	$43.54	$45.54	$44.57
Total	$48.66	$97.12	$47.96
3. Percent of Weeks Worked			
White	52.8	34.3	44.7
Non-white	49.7	36.0	41.5
Total	51.0	35.5	42.6
4. Average Hours of Work Per Week Worked			
White	31.8	32.7	32.1
Non-white	26.6	31.6	29.2
Total	28.9	32.0	30.3
5. Average Hourly Rate[3]			
White	$1.79	$1.56	$1.72
Non-white	$1.64	$1.44	$1.53
Total	$1.72	$1.47	$1.60
6. Number of Terminees Reporting			
White	264	201	467
Non-white	350	499	853
Total	619	704	1,323

[1] Data from study conducted by Dunlap and Associated, Inc.

[2] During post-program period covered by the study, terminees in other manpower programs, military service, or schools are treated as having no earnings.

[3] Based on different computer tabulation than the other items.

TABLE E.2

CHARACTERISTICS OF TERMINEES FROM JOB CORPS IN CITIES OF 1,000,000 OR MORE POPULATION USED FOR ANALYSIS *

Race of Terminee	Sex of Terminee		
	Male	Female	Total
1. Average Post-Job Corps Earnings Per Week, All Terminees			
White	$50.31	$43.56	$49.40
Non-white	$50.31	$26.60	$42.65
Total	$50.31	$28.73	$43.31
2. Average Pre-Job Corps Earnings Per Week, All Terminees			
White	$28.17	$ 5.06	$23.19
Non-white	$37.69	$34.90	$36.58
Total	$36.90	$33.37	$35.58
3. Average Post-Job Corps Earnings Per Week, Terminees With Any Work			
White	$62.89	$65.33	$63.44
Non-white	$62.29	$43.60	$57.31
Total	$62.33	$46.31	$57.98
4. Average Pre-Job Corps Earnings Per Week, Terminees With Any Work			
White	$37.56	$10.06	$34.77
Non-white	$56.54	$53.81	$55.27
Total	$54.79	$52.06	$53.60
5. Percent of Terminees With Any Work, Post-Job Corps			
White	80.0	66.7	77.3
Non-white	80.8	61.0	74.4
Total	80.7	62.0	74.7
6. Percent of Terminees With Any Work, Pre-Job Corps			
White	75.0	50.0	66.7
Non-white	66.7	64.9	66.2
Total	67.4	64.1	66.4

* Data from study conducted by Louis Harris and Associates, Inc.

TABLE E.2 (Cont)

7. Number of Terminees Reporting, Post-Job Corps			
White	101	46	147
Non-white	379	185	564
Total	480	231	711
8. Number of Terminees Reporting, Pre-Job Corps			
White	61	36	97
Non-white	238	147	385
Total	299	183	482

TABLE E.3

CHARACTERISTICS OF TERMINEES FROM JOB CORPS IN CITIES OF 250,000 - 999,999 POPULATION USED FOR ANALYSIS

Race of Terminee	Sex of Terminee		
	Male	Female	Total
1. Average Post-Job Corps Earnings Per Week, All Terminees			
White	$47.71	$26.10	$40.94
Non-white	$46.65	$27.58	$40.40
Total	$46.88	$27.29	$40.52
2. Average Pre-Job Corps Earnings Per Week, All Terminees			
White	$29.54	$17.76	$25.17
Non-white	$24.60	$12.48	$19.96
Total	$25.60	$13.51	$21.02
3. Average Post-Job Corps Earnings Per Week, Terminees With Any Work			
White	$56.04	$38.73	$51.44
Non-white	$55.96	$41.81	$52.02
Total	$55.98	$41.21	$51.90
4. Average Pre-Job Corps Earnings Per Week, Terminees With Any Work			
White	$40.04	$27.29	$35.90
Non-white	$40.10	$26.19	$35.58
Total	$40.08	$26.60	$35.65
5. Percent of Terminees Wtih Any Work, Post-Job Corps			
White	85.2	67.3	79.6
Non-white	83.4	66.0	77.7
Total	83.8	66.2	78.1
6. Percent of Terminees With Any Work, Pre-Job Corps			
White	73.8	63.9	70.1
Non-white	80.8	47.6	56.1
Total	63.9	50.8	58.9

TABLE E.3 (Cont)

7. Number of Terminees Reporting, Post-Job Corps			
White	15	7	22
Non-white	208	101	309
Total	223	108	331
8. Number of Terminees Reporting, Pre-Job Corps			
White	12	4	16
Non-white	136	74	210
Total	148	78	226
Source: Data from study conducted by Louis Harris and Associates, Inc.			

TABLE E.4

RELVARIANCE PER TERMINEE OF ANNUAL POST-PROGRAM EARNINGS BETWEEN NYC OUT-OF-SCHOOL TERMINEES WITHIN SITE

Race of Terminee	Sex of Terminee		
	Male	Female	Total
White	1.28	2.04	1.55
Non-white	1.57	1.81	1.71
Total	1.45	1.88	1.69

TABLE E.5
RELVARIANCE PER TERMINEE OF ANNUAL EARNINGS BETWEEN JOB CORPS TERMINEES WITHIN REGION, BY SIZE OF PLACE OF INTERVIEW

Race of Terminee	Sex of Terminee		
	Male	Female	Total
CITIES OF 1,000,000 OR MORE POPULATION			
Post-Job Corps Earnings			
White	1.34	0.84	1.06
Non-white	0.91	2.16	1.20
Total	0.93	1.95	1.19
Pre-Job Corps Earnings			
White	1.21	---	1.60
Non-white	2.16	1.69	2.03
Total	2.12	1.78	2 04
CITIES OF 250,000—999,999 POPULATION			
Post-Job Corps Earnings			
White	0.97	1.10	1.13
Non-white	0.90	1.73	1.12
Total	0.93	1.60	1.12
Pre-Job Corps Earnings			
White	1.51	3.65	2.09
Non-white	2.87	3.14	3.28
Total	2.53	3.57	2.98

TABLE E.6
RELVARIANCES PER TERMINEE IN AVERAGE ANNUAL POST-PROGRAM EARNINGS, BETWEEN TERMINEE WITHIN SITE, NYC OUT-OF-SCHOOL PROGRAM*

Race and Sex of Terminee	All Terminees		Terminees Who Worked One or More Weeks	
	Earnings per Week	Earnings per Week Worked	Earnings per Week	Earnings per Week Worked
All Terminees	1.69	0.88	0.89	0.60
White	1.55	0.72	0.92	0.52
Non-white	1.71	0.96	0.84	0.63
Male	1.45	0.93	0.96	0.71
Female	1.88	0.74	0.79	0.44
White:				
Male	1.28	0.65	0.91	0.53
Female	2.04	0.70	0.91	0.41
Non-white:				
Male	1.57	1.17	1.00	0.88
Female	1.81	0.76	0.72	0.44

*Based on data from survey by Dunlap and Associates, Inc.

TABLE E.7
PERCENTAGE DIFFERENCES BETWEEN COEFFICIENTS OF VARIATION FOR AVERAGE ANNUAL EARNINGS, AS ESTIMATED FROM SURVEY DATA AND DERIVED FROM GENERALIZED CURVES*

Race of Terminee	Sex of Terminee		
	Male	Female	Total
JOB CORPS TERMINEES			
(Cities of 1,000,000 or More Population)			
Post-Job Corps Earnings			
White	-13	14	-3
Non-white	4	-18	-4
Total	4	-15	-4
Pre-Job Corps Earnings			
White	32	NA	20
Non-white	-9	4	-6
Total	-8	9	-5
(Cities of 250,000-999,999 Population)			
Post-Job Corps Earnings			
White	3	16	--
Non-white	8	-9	1
Total	6	-5	--
Pre-Job Corps Earnings			
White	17	-15	2
Non-white	-11	- 1	-13
Total	- 9	- 9	- 9
NYC OUT-OF-SCHOOL TERMINEES			
White	4	- 8	1
Non-white	1	---	---
Total	2	- 3	- 2

* [(Derived - Estimated) ÷ Estimated] x100.

TABLE E.8

ESTIMATED AVERAGE RELVARIANCES PER SITE OF EARNINGS PER WEEK BETWEEN SITES WITHIN STRATUM, NYC OUT-OF-SCHOOL TERMINEE SURVEY DATA

Race and Sex of Terminee	Sites in Larger Urban Places		Sites in Smaller Urban Places	
	Earnings Per Week	Estimated Relvariance	Earnings Per Week	Estimated Relvariance
All terminees	$21.89	.015	$17.29	.11
White	$25.06	.19	$22.67	.084
Non-white	$20.55	.0065	$12.93	NA
Male	$25.84	.039	$22.83	.14
Female	$18.71	.013	$12.11	.11
White:				
Male	$29.71	.16	$28.16	.092
Female	$19.76	.18	$14.59	.18
Non-white:				
Male	$23.64	.022	$14.32	NA
Female	$18.51	.061	$ 9.29	NA

APPENDIX F

MATCHING OF CONTROL CASES WITH PROGRAM ENROLLEES

F.1 As background for determining the variables on which program enrollees and controls will be matched, it is useful to analyze the contribution of matching compared with regression analysis to the efficiency of an evaluation study.[1]

F.2 Consider first a study in which matching is carried out on only one variable. Suppose that with an "ideal" control population, i.e., one differing from the population sampled for the treatment group only in the experimental factor, the study observations can be expressed as:

Program Group: $y = \alpha + \beta x + e$

Control Group: $y = \alpha' + \beta x' + e'$

where y is the variable to be evaluated, x is the variable on which the matching is done $(\alpha - \alpha')$ is the true program effect, i.e., no unsuspected biases are present; and $\bar{X}$ and $\bar{X}'$, the means of X in the two populations, are equal. Assume that x and e are independently distributed and the $E(e) = E(e') = 0$, where E denotes the expected value or average of the indicated variable. Suppose that a sample of n matched pairs of observations is taken. The true program effect will be estimated by the difference in the mean value of y in the two groups

$$(\bar{y} - \bar{y}').$$

With this model, the variance of this estimate is

$$\sigma^2_{\bar{y} - \bar{y}'} = \frac{2}{n} \sigma^2 (1 - \rho^2)$$

---

[1] The discussion in this Appendix follows W.G. Cochran, "Matching in Analytical Studies," Amer. J. Pub. Health, 43(1953), 684-691.

where σ^2 is the variance of y, assumed to be the same for both populations, and ρ is the correlation of y and x. In general, if matching were done on multiple variables,

$$\sigma^2_{\bar{y} - \bar{y}'} = \frac{2}{n} \sigma^2 (1 - R^2)$$

where R is the multiple correlation between y and the set of variables used for matching. The impact of matching in reducing the variability of the estimated true program effect is shown in the following tabulation.

Measure of Variability	Value of Measure of Variability Compared to That With No Matching, When R Is							
	.2	.3	.4	.5	.6	.7	.8	.9
Variance	.96	.91	.84	.75	.64	.51	.36	.19
Sampling error	.98	.95	.92	.87	.80	.71	.60	.44

This tabulation indicates that the value of adding additional variables to the matching process depends upon the increase in R, and suggests that covariates having only small correlations with the variable of interest are not likely to produce much gain in precision. Thus, one needs to ask not whether a covariate is correlated with the variable of interest but what the size of the correlation is. Further, the reductions in sampling errors do not go over 10 percent until R reaches 0.5. In earlier evaluation studies of manpower training programs examined, this level of multiple correlation was about the highest found and was achieved with the use of over 10 covariates. To attempt matching on this number of covariates would be quite time-consuming and costly for the screening and, more seriously, likely to result in fairly distorted control samples. It is a relatively easy task to add to a list of potential covariables. On the other hand, at present there is only limited knowledge as to the magnitude of the correlations of such covariables with program impact measures of interest.

F.3 To compare the effect of regression adjustments with that of matching, suppose simple random samples of n program cases and n controls were drawn and the program effect estimated by the difference in the adjusted means of the two samples, say,

$$(\bar{y}^* - \bar{y}'^*).$$

The variance of this estimate is, with the assumed model[2/]

[2/] This assumes that x is normally distributed, but is approximately true if x is not normal; see W.G. Cochran, Sampling Techniques (2nd Ed.), John Wiley & Sons, Inc., New York, 1963.

$$\sigma^2_{\bar{y}^* - \bar{y}'^*} = \frac{2}{n} \sigma^2 \; (1 - \rho^2) \left\{ 1 + \frac{1}{2(n-2)} \right\} .$$

The factor in brackets represents the increase in variance compared with matching. With multiple regression based on k variables,

$$\sigma^2_{\bar{y}^* - \bar{y}'} = \frac{2}{n} \sigma^2 \; (1-R^2) \left\{ 1 + \frac{k}{2n - k + 3} \right\} .$$

For n, say greater than 10k, the factor in brackets is close to unity and regression adjustment provides about the same precision as matching. This assumes that the correct regression function is used, otherwise a bias term must be added.

F.4 If the program cases are such that $\overline{X}$ and $\overline{X}'$ differ,

$$\sigma^2_{\bar{y}^* - \bar{y}'^*} = \frac{2}{n} \sigma^2 \; (1 - R^2) \left\{ 1 + \frac{1}{2(n-2)} + \frac{n D^2}{4(n-2)\sigma^2_x} \right\}$$

where $D = \overline{X} - \overline{X}'$. For n, say over 20, the increase in variance with regression analysis compared with matching is

$$1 + \frac{D^2}{4\sigma_x^2}$$

independent of sample size. For example, if $D = \sigma_x$, which implies fairly drastic selection operating on the x variable in the recruiting of enrollees for the program

$$1 + \frac{D^2}{4\sigma^2_x} = 1.25.$$

F.5 The principal conclusions for planning the study are that if the program participants represent a highly selective group compared to the target population, the estimates of program impact based on matched controls are likely to give a biased estimate of the program effects to be expected for the larger population.

The bias can be expected to be smaller with matched than with unmatched samples, but may be only slightly smaller. If the appropriate regression form is employed, regression analysis can achieve about the same precision as matching with considerably less difficulty in the field survey procedures and administration.